Spectrum Sensing Using Correlated Receiving Multiple Antennas in Cognitive Radios

Saeid Sedighi, Abbas Taherpour, *Member, IEEE*, and Josep Sala, *Senior Member, IEEE*

***Abstract*—In this paper, we address the problem of multi-antenna spectrum sensing in Cognitive Radios (CRs) by considering the correlation between the received channels at different antennas. First, we derive the optimum genie-aided detector which assumes perfect knowledge of the antenna correlation coefficients, Primary User (PU) signal power and noise variance. This is used as a benchmark for comparing with more practical detectors when some or all of these parameters are unknown to the Secondary User (SU). Two scenarios are considered: 1) the antenna correlation coefficients and PU signal power are unknown to the SU; 2) the antenna correlation coefficients, PU signal power and noise variance are unknown to the SU. To derive sub-optimum detectors for these two scenarios, we apply the Rao test, an asymptotically equivalent test to the Generalized Likelihood Ratio Test (GLRT) that does not require the Maximum Likelihood (ML) estimates of unknown parameters. Additionally, we calculate analytical approximations to the detection and false-alarm probabilities of the proposed detectors and verify them with Monte-Carlo simulations. The simulation results show that these new detectors outperform several recently proposed detectors for CR using multiple antennas.**

***Index Terms*—Cognitive radio, spectrum sensing, multiple antennas, Rao test, antenna correlations, Fisher information matrix, noise variance mismatch, antenna array.**

I. Introduction

THE identification of spectrum holes by SUs constitutes a major requirement at the physical layer of CR networks, where spectrum sensing techniques are sought to achieve a sufficiently reliable detection probability over the shortest possible sensing time. So far, several different methods have been proposed for spectrum sensing [1], [2]. The Energy Detector (ED) is a popular method to detect an unknown signal in additive white noise [3], [4]. Unfortunately, the performance of the ED is susceptible to errors in the noise variance and it has been shown that to achieve a desired probability of detection under noise, or in more general terms, model uncertainties, the Signal-to-Noise Ratio (SNR) must be above a certain threshold [5] (SNR wall).

Manuscript received January 23, 2013; revised May 1 and July 12, 2013; accepted August 27, 2013. The associate editor coordinating the review of this paper and approving it for publication was C.-P. Li.

S. Sedighi and A. Taherpour are with the Dept. of Electrical Engineering, Imam Khomeini International University (IKIU), Qazvin, Iran (e-mail: {s.sedighi, taherpour}@ikiu.ac.ir).

J. Sala is with the Dept. of Signal Theory and Communications, Technical University of Catalonia - Barcelona Tech (UPC), Barcelona, Spain (e-mail: josep.sala@upc.edu).

This work was partially supported by projects AGAUR - 2009SGR1236, TEC2010-21245-C02-01 (DYNACS), CONSOLIDER INGENIO CSD2008-00010 (COMONSENS) and CENIT CEN-20101019 (THOFU) financed by the Catalan and Spanish Governments.

Digital Object Identifier 10.1109/TWC.2013.100213.130158

An efficient strategy to increase the reliability of spectrum sensing is to use cooperative sensing, in which information from multiple spatially distributed SUs is incorporated for the detection of the PU [6]. Recently, some research works have been conducted in that field [7]–[11]. [7] studies optimization of cooperative spectrum sensing with an improved ED in each SU over imperfect reporting channels. Cooperative spectrum sensing for a CR mesh network is considered in [8]. A linear cooperative sensing framework based on the combination of the observed energies by different SUs is proposed in [9]. In [10], [11], the authors propose a selective-relay based cooperative spectrum sensing scheme without a dedicated reporting channel, which is able to control and reduce the interference from SUs to the PU.

Using multiple antennas at the SU receiver is one possible approach to overcome noise uncertainty and also to improve the performance of spectrum sensing by exploiting available observations in the spatial domain. Another advantage of using multiple antennas at the SU receiver is not having to employ a (probably imperfect) reporting channel like that needed for cooperative spectrum sensing. Multiantenna techniques have been used by different authors for spectrum sensing [12]–[20]. [12] considers a blind spectrum sensing approach where the empirical characteristic function of the observed sample vectors from multiple antennas is used in the formulation of the statistical test. In [13], the authors derive the optimum Neyman-Pearson (NP) and sub-optimum GLRT-based multiantenna detectors of an Orthogonal Frequency Division Multiplexing (OFDM) signal with a cyclic prefix of known length. Multiantenna spectrum sensing in frequency selective channels is addressed in [14]. In [15]–[18], the authors derive the GLRT detectors of spatial rank-one PU signals robust to noise variance uncertainty. Finally, the GLRT eigenvalue-based detectors for multiantenna spectrum sensing are proposed in [19], [20] for PU signals with spatial rank larger than one.

Some works, such as [21], [22], show that, in general, stronger antenna correlations appear as the beamwidth of the received signal decreases. In CR networks, signals from far PUs arrive at the SU base station within a small beamwidth, which results in a high correlation between the channel gains of different antennas. As an example, we may refer to IEEE 802.22 Wireless Regional Area Network (WRAN) systems, a standard that defines the use of white spaces in the TV frequency spectrum [23], [24]. However, most research works on multiantenna spectrum sensing have assumed unknown deterministic channel gains at each antenna and have not considered statistical correlations between the antenna channel

gains. Multiantenna spectrum sensing considering statistical correlations between the antenna channel gains has recently been studied in some works [25]–[29]. The performance of the multiantenna ED over correlated Rayleigh and Nakagamim faded channels is derived in [25] and [26], respectively. In [27], the authors propose a weighted ED for multiantenna spectrum sensing with correlated receiving antennas when the antenna correlation coefficients, PU signal power and noise variance are known to the SU; Moreover, the authors derive a GLRT-based detector by using an approximation in the low-SNR regime when only the noise variance is unknown to the SU. In [29], the authors propose a detector that uses the weighted linear combination of the estimates of the cross correlation among the signals received at different antennas when the noise variance is unknown but the PU signal power and antenna correlation coefficients are known to the SU. Nevertheless, as shown in [25] and [26], the performance of the ED degrades significantly under the correlated receiving antennas model. Moreover, the proposed detectors in [27] and [29] have been derived assuming that the PU signal power and the antenna correlation coefficients are known to the SU, which, although useful for benchmarking, is not realistic in practice.

In this paper, we investigate multiantenna spectrum sensing by considering the correlation between the received channel gains at different antennas in the presence of Additive White Gaussian Noise (AWGN). First, for benchmarking purposes, we derive the optimal NP detector for the case when the SU has knowledge of the antenna correlation coefficients, PU signal power and noise variance. In practice, all of these parameters will be unknown to the SU. Hence, and for the sake of comparison, we consider two composite hypothesis scenarios: 1) the SU has no knowledge about the antenna correlation coefficients and the PU signal power; 2) additionally, the SU has no knowledge about the noise variance either. The GLRT is a useful scheme for such composite hypotheses, but, it needs to compute the ML parameter estimates. We show that for these two composite hypothesis scenarios, obtaining the ML estimates leads to an optimization problem for which a closed-form solution does not appear to exist. Thus, for deriving their corresponding sub-optimum detectors, we resort to the Rao test, which has the same asymptotic performance as the GLRT and does not require any parameter estimates. We evaluate the performance of the proposed detectors through analysis and simulation, and show they outperform several recently-proposed detectors for CR using multiple antennas.

Organization: Section II describes the system model and basic assumptions on the PU signal, channel gains and noise. In Section III, we derive the optimal NP detector. In Section IV, we present the new sub-optimum detectors based on the Rao test for the two mentioned scenarios. In Section V, the performance of the proposed detectors is evaluated analytically. Simulation results are shown and discussed in Section VI. Finally, conclusions are drawn in Section VII.

Notation: Lightface letters denote scalars. Boldface lower- and upper-case letters denote column vectors and matrices, respectively. $[\mathbf{A}]_{k,l}$ and x_i stand for the entries of matrix $\mathbf{A}$ and vector $\mathbf{x}$, respectively. The trace, determinant, inverse and Moore-Penrose pseudoinverse of matrix $\mathbf{A}$ are $\mathrm{tr}(\mathbf{A})$, $\det(\mathbf{A})$, $\mathbf{A}^{-1}$, $\mathbf{A}^{\dagger}$, respectively. $\mathrm{vec}[\mathbf{A}]$ is the column-wise vectorization of matrix $\mathbf{A}$. The $M \times M$ identity matrix is $\mathbf{I}_M$. Superscripts $*$, T and H are the complex conjugate, transpose and Hermitian (conjugate transpose), respectively. Letter j is the imaginary unit, i.e., $j^2 = -1$. The real and imaginary parts of x are $\mathrm{Re}\{x\}$ and $\mathrm{Im}\{x\}$, respectively. $\delta[\cdot]$ is Kronecker's delta. $\mathbf{\Lambda} = \mathrm{diag}\{x_1, \ldots, x_M\}$ is a diagonal matrix of entries $x_1, \ldots, x_M$. $\mathrm{E}\{\cdot\}$ is the statistical expectation. $\mathcal{CN}(\mathbf{m}, \mathbf{P})$ denotes the circular complex Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{P}$.

II. SYSTEM MODEL

A. Basic Assumptions and Problem Statement

We assume a single-antenna PU signal $x_P(t) = \mathrm{Re}[s(t)e^{j2\pi f_c t}]$ at the carrier frequency f_c, with $s(t) = i_P(t) + jq_P(t)$ its equivalent complex baseband signal of bandwidth B_P. We also assume, following related works [25], [27], [29], that the signal is low-rate (low bandwidth) so that a flat-fading model constitutes a close approximation to reality. For an M-antenna SU receiver, let $z_m(t) = x_{P,m}(t) + n_m(t)$ denote the signal at the m-th antenna, with $x_{P,m}(t)$ the received noiseless pass-band PU signal and $n_m(t)$ the corresponding additive white Gaussian noise process of power spectral density $\frac{N_0}{2}$ over the band of interest, with $n_m(t)$ independent and thus not correlated over different antennas. Each $z_m(t)$ is pass-band filtered at radio-frequency (with impulse response $h_{RF}(t)$), downconverted to baseband and sampled at the frequency $f_s = \frac{1}{T_s}$ to generate L complex temporal samples per antenna according to $y_m[i] = y_m(iT_s)$ with $y_m(t) = h_S(t) * ((z_m(t) * h_{RF}(t)) \cdot e^{-j2\pi f_c t}))$ and $h_S(t)$ the sampling filter impulse response.

Now, using $M \times 1$ vectors for stacking samples from different antennas at the i-th time instant, a discrete equivalent channel model can be compactly expressed in baseband as follows,

$$\mathbf{y}_i \quad = \quad \mathbf{h}_i s_i + \mathbf{n}_i, \tag{1}$$

with,

1) $\mathbf{h}_i \doteq [h_1[i], h_2[i], \cdots, h_M[i]]^T \in \mathbb{C}^M$ the equivalent random stationary complex channel gain vector at time iT_s modeling flat-fading. Its spatial correlation matrix is $\mathbf{C} = \mathrm{E}\{\mathbf{h}_i \mathbf{h}_i^H\}$.
2) $\mathbf{n}_i \doteq [n_1[i], n_2[i], \cdots, n_M[i]]^T \in \mathbb{C}^M$ the equivalent zero-mean i.i.d. circular complex baseband noise vector at time iT_s, with components $n_m[i]$ uncorrelated both in the spatial and temporal dimensions and of power σ^2. That is: $\mathrm{E}\{n_m[i] n^*_{m'}[i']\} = \sigma^2 \delta[m - m'] \delta[i - i']$.
3) $s_i \doteq s(iT_s) \in \mathbb{C}$ the samples of the transmitted baseband PU signal, of power $\mathrm{E}\{|s_i|^2\} = \mathcal{E}_s$.
4) $\{\mathbf{h}_i\}_{i=1}^L$, $\{\mathbf{n}_i\}_{i=1}^L$ and $\{s_i\}_{i=1}^L$ mutually independent and zero-mean sequences, and thus, mutually uncorrelated. In the following, any of $\mathbf{h}_i$, $\mathbf{n}_i$ or s_i will be referred to as time samples.

We make the following additional assumptions concerning the signal model,

1) We consider the detection of weak signals where long sensing times are required. Thus, in fast time-varying scenarios as vehicular communications, the channel may

undergo significative changes within the sensing window (on the order of several channel coherence times) [30] so that detectors based on the constant channel assumption may be severely impaired.

2) As the goal of this paper is the analysis of simple well-performing sub-optimum detectors which rely on spatial processing alone, no specific model for the temporal correlation of either $\mathbf{h}_i$ or of s_i is adopted. This approach, successfully applied in other works [16], [25], [27], [31], greatly simplifies the complexity of the detectors and their associated mathematical derivations and consists of taking the sequences $\mathbf{h}_i$ and s_i both uncorrelated in time (each with uniform power spectral density).
3) It is also our goal to derive a detector that tests for a suitable feature of the received signal: in our case, the specific spatial correlation pattern associated with the considered scenario (under the model described later in this section). Thus, we resort to the widely applied Gauss-in-Gauss model in the field of array processing, where the signal of interest at the receiver input, i.e. $\mathbf{h}_i s_i$, is considered Gaussian distributed [32], [33]. This approach leads to mathematically tractable and intuitively appealing detection and estimation schemes, usually operating on sample correlations of the received data, and which perform well in the low-SNR regime of interest [34]–[37]. In this context, the resulting detector is equivalent to testing for a particular correlation structure [31]. It is true that a detector based on the true Probability Density Function (PDF) of the signal would have better detection performance for a sufficiently long observation window (but possibly much more involved in terms of computational complexity). Nevertheless, as the derived detector operates on sample correlations, for a sufficiently long observation window which is definitely required in the low-SNR regime, those sample correlations tend to have an asymptotic Gaussian distribution for a broad range of possible PDF's of the signal of interest (those with finite fourth-order moments) [38]–[40]. We should note that for the particular case when s_i is constant modulus and $\mathbf{h}_i$ is Gaussian distributed, the random variable $\mathbf{h}_i s_i$ is also Gaussian distributed.
4) We assume that both the transmitter and the receivers use the same polarization. Hence, possible cross-polarization effects during propagation are not exploited by the detector. That would require a specific statistical model and falls out of the scope of this paper.

We denote the hypotheses of the presence and absence of the PU signal by $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively. The hypothesis testing problem at i-th time instant is then given by,

$$\mathcal{H}_0 : \{\mathbf{y}_i\}_{i=1}^{L} = \{\mathbf{n}_i\}_{i=1}^{L} \ , \ \mathcal{H}_1 : \{\mathbf{y}_i\}_{i=1}^{L} = \{\mathbf{h}_i s_i + \mathbf{n}_i\}_{i=1}^{L}. \tag{2}$$

Under the Gauss-in-Gauss model and when all signals are zero-mean, the PDF of the received data only depends on their correlations under each hypothesis $\mathcal{H}_\nu \in \{\mathcal{H}_0, \mathcal{H}_1\}$. For simplicity of notation, from now on we take $s_i = 0$ under $\mathcal{H}_0$, so that the complete set of correlations can be expressed from the autocorrelation function of the stationary vector sequence $\mathbf{y}_i$,

$$\begin{aligned} &\mathbf{R}_y[m|\mathcal{H}_\nu] = \mathrm{E}\{\mathbf{y}_{i+m}\mathbf{y}_i^H|\mathcal{H}_\nu\} \\ &= \mathrm{E}\{\mathbf{h}_{i+m}\mathbf{h}_i^H\} \cdot \mathrm{E}\{s_{i+m}s_i^*|\mathcal{H}_\nu\} + \mathrm{E}\{\mathbf{n}_{i+m}\mathbf{n}_i^H\} \\ &= \mathbf{C}[m] \cdot R_s[m|\mathcal{H}_\nu] + \sigma^2\delta[m] \cdot \mathbf{I}_M, \end{aligned} \tag{3}$$

with $\mathbf{C}[m]$ and $R_s[m|\mathcal{H}_\nu]$ implicitly defined. According to assumption no. 2 in the signal model, the channel and PU signal sequences are uncorrelated in time, with $\mathbf{C}[m] = \mathbf{C}\delta[m]$ and $R_s[m|\mathcal{H}_\nu] = \nu \cdot \mathcal{E}_s\delta[m]$, respectively. Hence, $\mathbf{R}_y[m|\mathcal{H}_\nu] = \mathbf{\Gamma}_\nu\delta[m]$, with $\mathbf{\Gamma}_\nu \doteq \nu\mathbf{\Sigma} + \sigma^2\mathbf{I}_M$ and $\mathbf{\Sigma} \doteq \mathcal{E}_s\mathbf{C}$. Let us define the observation matrix $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L] \in \mathbb{C}^{M\times L}$ containing the L time samples at the M antennas and vector $\mathbf{y} = \mathrm{vec}[\mathbf{Y}] \in \mathcal{C}^{ML\times 1}$. Now, defining the correlation matrix of $\mathbf{y}$: $\mathbf{R}_{y|\mathcal{H}_\nu} \doteq \mathrm{E}\{\mathbf{y}\mathbf{y}^H|\mathcal{H}_\nu\}$, we have that it is a block matrix whose $M \times M$ block in the (i, i') position is given by $\mathbf{R}_y[i'-i|\mathcal{H}_\nu] = \mathbf{\Gamma}_\nu\delta[i'-i]$. Thus, $\mathbf{R}_y$ is block-diagonal with equal blocks $\mathbf{\Gamma}_\nu$, or, $\mathbf{R}_{y|\mathcal{H}_\nu} = \mathbf{I}_L \otimes \mathbf{\Gamma}_\nu$, with $\otimes$ Kronecker's product. Then, using the properties of Kronecker's product, the PDF of $\mathbf{y}$ becomes,

$$\begin{aligned} f(\mathbf{y}|\mathcal{H}_\nu, \mathbf{\Gamma}_\nu) &= \frac{1}{\pi^{ML}\det(\mathbf{R}_{y|\mathcal{H}_\nu})} \exp\left(-\mathbf{y}^H\mathbf{R}_{y|\mathcal{H}_\nu}^{-1}\mathbf{y}\right) \\ &= \frac{\exp\left(-\mathrm{tr}\left(\mathbf{Y}\mathbf{Y}^H\mathbf{\Gamma}_\nu^{-1}\right)\right)}{(\pi^M\det(\mathbf{\Gamma}_\nu))^L}, \end{aligned} \tag{4}$$

where the multiple antenna PU detection problem can be expressed as the binary hypothesis test,

$$\begin{cases} \mathcal{H}_0 : & \mathbf{y} \sim \mathcal{CN}(\mathbf{0}_{M\cdot L}, \mathbf{I}_L \otimes (\sigma^2\mathbf{I}_M)), \\ \mathcal{H}_1 : & \mathbf{y} \sim \mathcal{CN}(\mathbf{0}_{M\cdot L}, \mathbf{I}_L \otimes (\mathbf{\Sigma} + \sigma^2\mathbf{I}_M)). \end{cases} \tag{5}$$

Although we adopt the Gaussian model in (5) for the sake of mathematical tractability and detector complexity, posterior performance evaluation shows that for signals s_i with more complex distributions (i.e. QAM) or with non-white spectrum (i.e. $R_s[m|\mathcal{H}_1] \neq \mathcal{E}_s\delta[m]$), the derived sub-optimum detectors still operate efficiently as the statistical model in (5) is capturing the essential features for detection. As will be seen, detector complexity after computing the sample correlations only requires a linear combination of the modulus-squared traces of K distinct $M \times M$ matrix products.

B. Antenna Correlation Model

Following [21], if the angle of arrival is uniformly distributed in the interval $[\phi-\Delta, \phi+\Delta]$, with ϕ the incident angle with respect to a reference frame and 2Δ the beamwidth of the PU signal received at the SU base station, then the correlation between the channel gains $h_n[i]$ and $h_m[i]$ of the n-th and m-th antennas (the component of the correlation matrix $\mathbf{C}$), is given by,

$$\mathrm{E}\{h_n[i]h_m^*[i]\} = [\mathbf{C}]_{n,m} = C_R^{n,m} + jC_I^{n,m}, \tag{6}$$

where, as in [29], the following correlation model from [21] is adopted,

$$\begin{aligned} & C_R^{n,m} \\ &\triangleq \mathrm{E}\left\{\mathrm{Re}\{h_n[i]\}\,\mathrm{Re}\{h_m[i]\}\right\} + \mathrm{E}\left\{\mathrm{Im}\{h_n[i]\}\,\mathrm{Im}\{h_m[i]\}\right\} \\ &= J_0\left(\frac{2\pi D_{n,m}}{\lambda_c}\right) \\ &\quad + 2\sum_{d=1}^{\infty} J_{2d}\left(\frac{2\pi D_{n,m}}{\lambda_c}\right)\cos(2d\phi_{n,m})\frac{\sin(2d\Delta)}{2d\Delta}, \end{aligned} \tag{7}$$

$$\begin{aligned} & C_I^{n,m} \\ &\triangleq \mathrm{E}\left\{\mathrm{Im}\{h_n[i]\}\,\mathrm{Re}\{h_m[i]\}\right\} - \mathrm{E}\left\{\mathrm{Re}\{h_n[i]\}\,\mathrm{Im}\{h_m[i]\}\right\} \\ &= 2\sum_{d=1}^{\infty} J_{2d+1}\left(\frac{2\pi D_{n,m}}{\lambda_c}\right)\sin\left((2d+1)\phi_{n,m}\right)\frac{\sin(2d\Delta)}{2d\Delta}, \end{aligned} \tag{8}$$

with $D_{n,m}$ the antenna spacing between the n-th and m-th antennas, $\phi_{n,m}$ the angle between the incoming wavefront and the segment linking the two antennas, λ_c the wavelength of the PU signal and $J_d(.)$ the Bessel function of the first kind and order d. For 1-D or 2-D regular arrays, several entries in matrix $\mathbf{C}$ feature the same value, as different pairs of antennas within the array have the same inter-element distance and angle pairs $(D_{n,m}, \phi_{n,m}) = (D_k, \phi_k)$ [32]. Therefore, if there are $K' = 2K - 1$ distinct entries in matrix $\mathbf{C}$, each corresponding to antennas with one of the K' distinct (D_k, ϕ_k) pairs, the following decomposition of a Hermitian matrix $\mathbf{\Sigma}$ can be applied to $\mathbf{C}$, where $\mathbf{\Sigma} = \mathcal{E}_s\mathbf{C}$ as defined,

$$\mathbf{\Sigma} = \varphi_1\mathbf{T}_1 + \sum_{k=2}^{K}\varphi_k\mathbf{T}_k + \sum_{k=2}^{K}\varphi_k^*\mathbf{T}_k^H, \tag{9}$$

with $\boldsymbol{\varphi} \triangleq [\varphi_1, \ldots, \varphi_K]$ the generating vector containing the distinct entries of matrix $\mathbf{\Sigma}$ and $\{\mathbf{T}_1 = \mathbf{I}_M, \{\mathbf{T}_k, \mathbf{T}_k^H\}_{k=2}^K\}$ an orthogonal basis of the space of patterned Hermitian matrices $\mathbf{\Sigma}$, with components defined as: $[\mathbf{T}_k]_{n,m} = \delta\left[D_{n,m} - D_k\right]\delta\left[\phi_{n,m} - \phi_k\right]$.

Special case 1. For the linear equispaced antenna array (the most typical 1-D antenna array [41]), all $\phi_{n,m}$ are equal and $\mathbf{\Sigma}$ becomes a Hermitian Toeplitz covariance matrix, which can be described by its first row $\boldsymbol{\beta} \triangleq [\beta_1, \ldots, \beta_M]$ as,

$$\mathbf{\Sigma} = \beta_1\mathbf{Q}_1 + \sum_{k=2}^{M}\beta_k\mathbf{Q}_k + \sum_{k=2}^{M}\beta_k^*\mathbf{Q}_k^H, \tag{10}$$

where $\{\mathbf{Q}_k\}_{k=1}^M$ is the orthogonal basis of the space of Hermitian Toeplitz matrices associated with the sub-diagonals, of components expressed as: $[\mathbf{Q}_k]_{n,m} = \delta\left[(n-m) - (k-1)\right]$.

Special case 2. For the uniform circular antenna array (a typical 2-D antenna array [41]), the angles $\phi_{n,m}$ take different values. Defining $\mathbf{\Sigma}$ is more complex and we only show the specific case $M = 4$ (square array), labeling the antennas from 1 to $M = 4$ counter-clockwise. Then, we have $K' = 9 = 2 \times 5 - 1$ distinct (D_k, ϕ_k) pairs that lead to the following pattern for $\mathbf{\Sigma}$,

$$\begin{aligned} \mathbf{\Sigma} &= \begin{bmatrix} \varrho_1 & \varrho_2 & \varrho_4 & \varrho_3 \\ \varrho_2^* & \varrho_1 & \varrho_3 & \varrho_5 \\ \varrho_4^* & \varrho_3^* & \varrho_1 & \varrho_2^* \\ \varrho_3^* & \varrho_5^* & \varrho_2 & \varrho_1 \end{bmatrix} \\ &= \varrho_1\mathbf{G}_1 + \sum_{k=2}^{K=5}\varrho_k\mathbf{G}_k + \sum_{k=2}^{K=5}\varrho_k^*\mathbf{G}_k^H, \end{aligned} \tag{11}$$

with $\{\mathbf{G}_k\}_{k=1}^5$ implicitly defined.

III. Optimum Detector

In this section, we compute the NP detector corresponding to the case in which the SU has complete knowledge of the covariance matrix $\mathbf{\Sigma}$ and of the noise variance σ^2. From (5), the PDF of $\mathbf{y}$ under hypothesis $\mathcal{H}_0$ is given by,

$$f(\mathbf{y}|\mathcal{H}_0, \sigma^2) = \frac{1}{(\pi\sigma^2)^{ML}}\exp\left\{\frac{-L\mathrm{tr}\left(\mathbf{R}\right)}{\sigma^2}\right\}, \tag{12}$$

with $\mathbf{R} = \frac{1}{L}\mathbf{Y}\mathbf{Y}^H$. Taking the logarithm in (12) and defining $\mathcal{L}_0(\mathbf{y}) \triangleq \ln f(\mathbf{y}|\mathcal{H}_0, \sigma^2)$, we have,

$$\mathcal{L}_0(\mathbf{y}) = -\frac{L\mathrm{tr}\left(\mathbf{R}\right)}{\sigma^2} - ML\ln\pi - LM\ln\sigma^2. \tag{13}$$

Similarly, from (5), the PDF of $\mathbf{y}$ under hypothesis $\mathcal{H}_1$ is,

$$f(\mathbf{y}|\mathcal{H}_1, \mathbf{\Sigma}, \sigma^2) = \frac{\exp\left\{-L\mathrm{tr}\left(\mathbf{R}(\mathbf{\Sigma} + \sigma^2\mathbf{I}_M)^{-1}\right)\right\}}{(\pi^M\det(\mathbf{\Sigma} + \sigma^2\mathbf{I}_M))^L}. \tag{14}$$

Taking the logarithm and defining $\mathcal{L}_1(\mathbf{y}) \triangleq \ln f(\mathbf{y}|\mathcal{H}_1, \mathbf{\Sigma}, \sigma^2)$, we get,

$$\begin{aligned} \mathcal{L}_1(\mathbf{y}) &= -L\mathrm{tr}\left(\mathbf{R}(\mathbf{\Sigma} + \sigma^2\mathbf{I}_M)^{-1}\right) - ML\ln\pi \\ &\quad -L\ln\det(\mathbf{\Sigma} + \sigma^2\mathbf{I}_M). \end{aligned} \tag{15}$$

Now, from (13) and (15), the Log-Likelihood Ratio (LLR) function can be calculated as,

$$\begin{aligned} \mathrm{LLR} &= \mathcal{L}_1(\mathbf{y}) - \mathcal{L}_0(\mathbf{y}) \\ &= -L\mathrm{tr}\left(\mathbf{R}(\mathbf{\Sigma} + \sigma^2\mathbf{I}_M)^{-1}\right) + \frac{L\mathrm{tr}\left(\mathbf{R}\right)}{\sigma^2} \\ &\quad - L\ln\det(\mathbf{\Sigma} + \sigma^2\mathbf{I}_M) + LM\ln\sigma^2. \end{aligned} \tag{16}$$

By using the matrix inversion lemma, we achieve,

$$(\mathbf{\Sigma} + \sigma^2\mathbf{I}_M)^{-1} = \sigma^{-2}\mathbf{I}_M - \sigma^{-4}(\mathbf{\Sigma}^{-1} + \sigma^{-2}\mathbf{I}_M)^{-1}. \tag{17}$$

By substituting (17) in (16), we have,

$$\begin{aligned} \mathrm{LLR} &= \frac{L\mathrm{tr}\left(\mathbf{R}(\mathbf{\Sigma}^{-1} + \sigma^{-2}\mathbf{I}_M)^{-1}\right)}{\sigma^4} - L\ln\det(\mathbf{\Sigma} + \sigma^2\mathbf{I}_M) \\ &\quad + LM\ln\sigma^2. \end{aligned} \tag{18}$$

Comparing the LLR in (18) with the threshold τ, the optimum detector according to the NP criterion can be expressed as,

$$T_{\mathrm{opt}} = L\mathrm{tr}\left(\mathbf{R}(\mathbf{\Sigma}^{-1} + \sigma^{-2}\mathbf{I}_M)^{-1}\right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{19}$$

where $\eta = \sigma^4(\tau + L\ln\det(\mathbf{\Sigma} + \sigma^2\mathbf{I}_M) - LM\ln\sigma^2)$.

We can exploit the Singular Value Decomposition (SVD) of the covariance matrix $\mathbf{\Sigma}$: $\mathbf{\Sigma} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, where the columns of $\mathbf{U}$ contain the eigenvectors of $\mathbf{\Sigma}$ and $\mathbf{\Lambda} =$

$\mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_M\}$ its corresponding eigenvalues. Then, (19) can be written as,

$$\begin{aligned} T_{\mathrm{opt}} =& L\mathrm{tr}\left(\mathbf{R}\mathbf{U}(\mathbf{\Lambda}^{-1}+\sigma^{-2}\mathbf{I}_M)^{-1}\mathbf{U}^H\right) \\ =& \mathrm{tr}\left(\mathbf{U}^H\mathbf{Y}\mathbf{Y}^H\mathbf{U}(\mathbf{\Lambda}^{-1}+\sigma^{-2}\mathbf{I}_M)^{-1}\right) \underset{<}{\overset{>}{}}_{\mathcal{H}_0}^{\mathcal{H}_1} \eta. \end{aligned} \quad (20)$$

Let $\mathbf{W} \triangleq \mathbf{U}^H\mathbf{Y}$. Then,

$$\begin{aligned} T_{\mathrm{opt}} =& \mathrm{tr}\left(\mathbf{W}\mathbf{W}^H(\mathbf{\Lambda}^{-1}+\sigma^{-2}\mathbf{I}_M)^{-1}\right) \underset{<}{\overset{>}{}}_{\mathcal{H}_0}^{\mathcal{H}_1} \eta \\ \Rightarrow \quad T_{\mathrm{opt}} =& \sum_{i=1}^{M} \frac{\sigma^2\lambda_i}{\sigma^2+\lambda_i}\|\mathbf{w}_i\|^2 \underset{<}{\overset{>}{}}_{\mathcal{H}_0}^{\mathcal{H}_1} \eta, \end{aligned} \quad (21)$$

where $\mathbf{w}_i^H \in \mathbb{C}^M$ is the i-th column of $\mathbf{W}^H$. In the optimum detector in (21), the observation matrix $\mathbf{Y}$ is first transformed to $\mathbf{W}$ under the linear transformation $\mathbf{U}$. Then, the weighted sum of the energies of the columns of $\mathbf{W}^H$ is calculated and compared with a given threshold.

IV. Proposed Sub-Optimum Detectors

In practical scenarios, $\mathbf{\Sigma}$ and/or σ^2 will be unknown to the SU. One useful solution for such composite hypothesis testing problems is the GLRT. The GLRT first obtains the ML estimates of unknown parameters under $\mathcal{H}_0$ and $\mathcal{H}_1$ and then forms the GLRT statistic as,

$$\mathrm{LR} \quad = \quad \frac{f(\mathbf{y}|\mathcal{H}_1, \widehat{\mathbf{\Theta}}_1)}{f(\mathbf{y}|\mathcal{H}_0, \widehat{\mathbf{\Theta}}_0)}, \quad (22)$$

with $\widehat{\mathbf{\Theta}}_\nu = \underset{\mathbf{\Theta}_\nu}{\arg\max}\ f(\mathbf{y}|\mathcal{H}_\nu, \mathbf{\Theta}_\nu)$ the ML estimates of unknown parameters under $\mathcal{H}_\nu$. However, as most often happens in realistic scenarios, analytical expressions for these ML estimates may be hard or impossible to obtain. Instead, we can use asymptotically equivalent tests to the GLRT that do not need the ML parameter estimates under $\mathcal{H}_1$ and which are easier to compute.

In this section, we consider two spectrum sensing scenarios: (1) $\mathbf{\Sigma}$ is unknown to the SU; (2) both $\mathbf{\Sigma}$ and σ^2 are unknown to the SU. Then, firstly, we show that for both scenarios it is not possible to obtain the ML parameter estimates under $\mathcal{H}_1$ in closed form. Secondly, we apply the Rao test as an asymptotically equivalent test to the GLRT which does not require any parameter estimates under $\mathcal{H}_1$ and we derive the corresponding sub-optimum detector for each scenario.

A. $\mathbf{\Sigma}$ unknown and σ^2 known

Using (9), the ML estimate of $\mathbf{\Sigma}$ under $\mathcal{H}_1$ can be obtained from the ML estimate of $\boldsymbol{\varphi}$ as,

$$\widehat{\boldsymbol{\varphi}} = \underset{\boldsymbol{\varphi}}{\arg\max}\ f(\mathbf{y}|\mathcal{H}_1, \mathbf{\Sigma}, \sigma^2). \quad (23)$$

Differentiating (15) with respect to φ_k for $1 \le k \le K$ and equating to zero, we get the equation,

$$\begin{aligned} \frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \varphi_k} =& -L\mathrm{tr}\left((\mathbf{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}\mathbf{T}_k\right) \\ &+ L\mathrm{tr}\left((\mathbf{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}\mathbf{T}_k(\mathbf{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}\mathbf{R}\right) = 0, \end{aligned} \quad (24)$$

which cannot be solved for φ_k analytically. Thus, it is not possible to obtain the ML estimate of $\boldsymbol{\varphi}$ in closed form. So, for this scenario, we propose to apply the Rao test rather than the GLRT.

The Rao Test (No Nuisance Parameters). *Let $\boldsymbol{\mu} \in \mathbb{R}^r$ be the set of parameters under $\mathcal{H}_1$ and $\mathcal{H}_0$. For a composite hypothesis test of the form,*

$$\mathcal{H}_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \quad , \quad \mathcal{H}_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \ , \quad (25)$$

the Rao test statistic is given by,

$$T_{R1} = \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \boldsymbol{\mu}}\right|_{\boldsymbol{\mu}=\boldsymbol{\mu}_0}^{T} \mathbf{I}^{-1}(\boldsymbol{\mu}_0) \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \boldsymbol{\mu}}\right|_{\boldsymbol{\mu}=\boldsymbol{\mu}_0}, \quad (26)$$

where $\mathbf{I}(\boldsymbol{\mu}_0)$ is the Fisher information matrix evaluated at $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ [33].

In this section, we have that $\boldsymbol{\mu} = [\varphi_1, \mathrm{Re}\{\varphi_2\}, \ldots, \mathrm{Re}\{\varphi_K\}, \mathrm{Im}\{\varphi_2\}, \ldots, \mathrm{Im}\{\varphi_K\}]^T \in \mathbb{R}^{2K-1}$ and $\boldsymbol{\mu}_0 = \mathbf{0}$ for the mentioned scenarios. Hence, using the Rao test yields the following detector,

$$\begin{aligned} T_{R1} =& \sum_{k=2}^{K} \frac{2L|\mathrm{tr}(\mathbf{T}_k\mathbf{R})|^2}{\sigma^4\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)^2\right)} \\ &+ LM\left(\frac{\mathrm{tr}(\mathbf{R})}{M\sigma^2}-1\right)^2 \underset{<}{\overset{>}{}}_{\mathcal{H}_0}^{\mathcal{H}_1} \eta. \end{aligned} \quad (27)$$

The application of the Rao test for this scenario and the derivation of the detector T_{R1} are detailed in Appendix A. Note that the first term in (27) operates on the projections of the off-diagonal elements of $\mathbf{R}$ onto $\mathbf{T}_k$, i.e., $\mathrm{tr}(\mathbf{T}_k\mathbf{R})$, which, asymptotically for a long sensing time, only depend on the second-order statistics of the signal of interest $\mathbf{h}_i s_i$. In fact, in accordance with the definition of $\mathbf{T}_k$, the term $t_k \triangleq \mathrm{tr}(\mathbf{T}_k\mathbf{R})/\sqrt{\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)^2\right)}$ is directly averaging those components of $\mathbf{R}$ whose expected value should be equal according to the spatial correlation model of the antenna gains. Finally, the first term in (27) is $\frac{2}{\sigma^4}$ times the combination of the energies $|t_k|^2$ of each projection: $\sum_{k=2}^{K}|t_k|^2$. The second term $\frac{\mathrm{tr}(\mathbf{R})}{M\sigma^2}-1$ utilizes the energy computation of the main diagonal elements of $\mathbf{R}$ and it only depends (asymptotically) on the power of the signal of interest as the noise power σ^2 gets canceled out. In summary, the detector in (27) is a test for a spatial correlation structure using the sample spatial correlation $\mathbf{R}$.

B. Both $\mathbf{\Sigma}$ and σ^2 unknown

We assume in this section that both σ^2 and $\mathbf{\Sigma}$ are unknown to the SU. By setting $\frac{\partial \mathcal{L}_0(\mathbf{y})}{\partial \sigma^2} = 0$, the ML estimate of σ^2 under $\mathcal{H}_0$ is obtained as,

$$\widehat{\sigma}_0^2 \quad = \quad \frac{1}{M}\mathrm{tr}(\mathbf{R}). \quad (28)$$

To obtain the ML estimates of σ^2 and $\mathbf{\Sigma}$ under $\mathcal{H}_1$, we need to solve the following system of equations with respect to $\boldsymbol{\varphi}$ and σ^2,

$$\begin{cases} \frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \boldsymbol{\varphi}} = \mathbf{0}, \\ \frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \sigma^2} = 0, \end{cases} \quad (29)$$

where $\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \varphi_k}$ for $1 \leq k \leq K$ appears in (24). Differentiating (15) with respect to σ^2, we get,

$$\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \sigma^2} = -L\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}\right) + L\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-2}\mathbf{R}\right). \quad (30)$$

It is easy to verify that the system of equations (29) cannot be solved in closed form with respect to $\boldsymbol{\varphi}$ and σ^2. Thus, for this scenario, we alternatively propose the Rao test rather than the GLRT.

The Rao Test (Nuisance Parameters). *Let $\boldsymbol{\theta} = [\boldsymbol{\mu}_r^T, \boldsymbol{\mu}_s^T]^T$, with $\boldsymbol{\mu}_r \in \mathbb{R}^r$ and $\boldsymbol{\mu}_s \in \mathbb{R}^s$, be the set of parameters under $\mathcal{H}_1$ and $\mathcal{H}_0$. For a composite hypothesis test of the form,*

$$\mathcal{H}_0 : \boldsymbol{\mu}_r = \boldsymbol{\mu}_{r_0}, \boldsymbol{\mu}_s \qquad , \qquad \mathcal{H}_1 : \boldsymbol{\mu}_r \neq \boldsymbol{\mu}_{r_0}, \boldsymbol{\mu}_s \ , \quad (31)$$

the Rao test statistic is given by,

$$T_{R2} = \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \boldsymbol{\mu}_r}\right|^T_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} [\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r} \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \boldsymbol{\mu}_r}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0}, \quad (32)$$

where $\widehat{\boldsymbol{\theta}}_0 \triangleq [\boldsymbol{\mu}_{r_0}^T, \widehat{\boldsymbol{\mu}}_{s_0}^T]^T$ is the ML estimate of $\boldsymbol{\theta}$ under $\mathcal{H}_0$ and,

$$[\mathbf{I}^{-1}(\boldsymbol{\theta})]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r} = \left(\mathbf{I}_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}(\boldsymbol{\theta}) - \mathbf{I}_{\boldsymbol{\mu}_r\boldsymbol{\mu}_s}(\boldsymbol{\theta})\mathbf{I}^{-1}_{\boldsymbol{\mu}_s\boldsymbol{\mu}_s}(\boldsymbol{\theta})\mathbf{I}_{\boldsymbol{\mu}_s\boldsymbol{\mu}_r}(\boldsymbol{\theta})\right)^{-1}. \quad (33)$$

The matrices in (33) are partitions of the Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$, defined as,

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{I}_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}(\boldsymbol{\theta}) & \mathbf{I}_{\boldsymbol{\mu}_r\boldsymbol{\mu}_s}(\boldsymbol{\theta}) \\ \mathbf{I}_{\boldsymbol{\mu}_s\boldsymbol{\mu}_r}(\boldsymbol{\theta}) & \mathbf{I}_{\boldsymbol{\mu}_s\boldsymbol{\mu}_s}(\boldsymbol{\theta}) \end{bmatrix}. \quad (34)$$

In this section, we have that $\boldsymbol{\mu}_r = [\varphi_1, \mathrm{Re}\{\varphi_2\}, \ldots, \mathrm{Re}\{\varphi_K\}, \mathrm{Im}\{\varphi_2\}, \ldots, \mathrm{Im}\{\varphi_K\}]^T \in \mathbb{R}^{2K-1}$, $\boldsymbol{\mu}_{r_0} = \mathbf{0}$ and $\mu_s = \sigma^2 \in \mathbb{R}$ for the mentioned scenarios. The application of the Rao test carried out in Appendix B demonstrates that the usual expression for the inverse $\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}$ yields a singular matrix for this scenario, which is a contradiction. Hence, we apply a modified Rao test, in which we use the Moore-Penrose pseudoinverse [42, Ch. 5.12] of $\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}$ instead of $[\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}$, to derive the corresponding sub-optimum detector. The application of the modified Rao test yields the following detector,

$$T_{R2} = \sum_{k=2}^{K} \frac{2L|\mathrm{tr}(\mathbf{T}_k\mathbf{R})|^2}{\widehat{\sigma}_0^4 \mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)^2\right)} \underset{<}{\overset{\geq}{\mathcal{H}_0}} \eta. \quad (35)$$

The application of the Rao test to this scenario and the derivation of T_{R2} are found in Appendix B. In comparison with (27), the second term does not appear as both the signal and the noise power are unknown: they are not identifiable from the diagonal of $\mathbf{R}$. Moreover, (35) and the first term in (27) coincide except for that, in (35), $\widehat{\sigma}$ substitutes the role of σ: the test becomes invariant to arbitrary scalings of the received data, a desirable feature under noise uncertainty.

V. Analytical Performance

In the following sub-sections, we evaluate the performance of the proposed detectors in terms of the detection and false-alarm probabilities P_d and P_fa, respectively, by deriving an approximation to their statistical distribution under $\mathcal{H}_1$ and $\mathcal{H}_0$. We assume that the proposed statistical model exactly matches the true distribution of the received data in order to capitalize on the theory of the Rao test. The validity of these results is verified in the simulations section.

A. Performance of the Optimum Detector

From (4) and (5), under hypothesis $\mathcal{H}_0$, all the elements of $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L] \in \mathbb{C}^{M\times L}$ are i.i.d. zero-mean complex Gaussian random variables with variance σ^2. Thus, under the linear transformation $\mathbf{W} = \mathbf{U}\mathbf{Y}$, with $\mathbf{U}\mathbf{U}^H = \mathbf{I}_M$, all the elements of $\mathbf{W}$ are also i.i.d. zero-mean complex Gaussian random variables with variance σ^2. Therefore, under $\mathcal{H}_0$, the random variable $\|\mathbf{w}_i\|^2$ is the summation of L i.i.d. random variables and, from the Central Limit Theorem (CLT) and L sufficiently large ($L > 15$), $\|\mathbf{w}_i\|^2$ is approximately Gaussian distributed [43] as $\|\mathbf{w}_i\|^2 \sim \mathcal{N}(L\sigma^2, L\sigma^4)$. Consequently, the decision statistic T_opt under $\mathcal{H}_0$ is a linear combination of independent Gaussian random variables, i.e., $\|\mathbf{w}_i\|^2$, and thus is Gaussian distributed as,

$$T_\mathrm{opt} \sim \mathcal{N}(\mu_{\mathrm{T}|\mathcal{H}_0}, \sigma^2_{\mathrm{T}|\mathcal{H}_0}), \quad (36)$$

where, defining $\mu_{\mathrm{T}|\mathcal{H}_0} = \mathrm{E}\{T_\mathrm{opt}|\mathcal{H}_0\}$, we have,

$$\mu_{\mathrm{T}|\mathcal{H}_0} = \sum_{i=1}^{M} \frac{\sigma^2\lambda_i}{\sigma^2+\lambda_i}\mathrm{E}\{\|\mathbf{w}_i\|^2|\mathcal{H}_0\} = \sum_{i=1}^{M}\frac{L\sigma^4\lambda_i}{\sigma^2+\lambda_i}, \quad (37)$$

and, setting $\sigma^2_{\mathrm{T}|\mathcal{H}_0} = \mathrm{E}\{T^2_\mathrm{opt}|\mathcal{H}_0\} - \mu^2_{\mathrm{T}|\mathcal{H}_0}$,

$$\begin{aligned}\sigma^2_{\mathrm{T}|\mathcal{H}_0} &= \sum_{i=1}^{M}\frac{\sigma^4\lambda_i^2}{(\sigma^2+\lambda_i)^2}\mathrm{E}\{\|\mathbf{w}_i\|^4|\mathcal{H}_0\} \\ &+2\sum_{i<k}\frac{\sigma^4\lambda_i\lambda_k}{(\sigma^2+\lambda_i)(\sigma^2+\lambda_k)}\mathrm{E}\{\|\mathbf{w}_i\|^2|\mathcal{H}_0\}\mathrm{E}\{\|\mathbf{w}_k\|^2|\mathcal{H}_0\} \\ &\quad -\sum_{i=1}^{M}\frac{L^2\sigma^8\lambda_i^2}{(\sigma^2+\lambda_i)^2} - 2\sum_{i<k}\frac{L^2\sigma^8\lambda_i\lambda_k}{(\sigma^2+\lambda_i)(\sigma^2+\lambda_k)} \\ &= \sum_{i=1}^{M}\frac{L\sigma^8\lambda_i^2}{(\sigma^2+\lambda_i)^2}. \end{aligned} \quad (38)$$

On the other hand, from (5), the columns of $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_L] \in \mathbb{C}^{M\times L}$ are distributed under $\mathcal{H}_1$ as independent zero-mean complex Gaussian vectors with covariance matrix $\boldsymbol{\Sigma} + \sigma^2\mathbf{I}_M$. Thus, under the linear transformation $\mathbf{W} = \mathbf{U}\mathbf{Y}$, where the columns of $\mathbf{U}$ contain the eigenvectors of $\boldsymbol{\Sigma}$, the columns of $\mathbf{W}$ are also independent zero-mean complex Gaussian distributed with covariance matrix $\mathbf{U}^H(\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)\mathbf{U} = \boldsymbol{\Lambda} + \sigma^2\mathbf{I}_M$. Therefore, under $\mathcal{H}_1$, the random variable $\|\mathbf{w}_i\|^2$ is the summation of L i.i.d. random variables and, according to the CLT, for a sufficiently large L, it will be approximately distributed as $\|\mathbf{w}_i\|^2 \sim \mathcal{N}\left(L(\lambda_i+\sigma^2), L(\lambda_i+\sigma^2)^2\right)$. Therefore, the decision statistic T_opt under $\mathcal{H}_1$ has the following Gaussian distribution,

$$T_\mathrm{opt} \sim \mathcal{N}(\mu_{\mathrm{T}|\mathcal{H}_1}, \sigma^2_{\mathrm{T}|\mathcal{H}_1}), \quad (39)$$

where, defining $\mu_{\mathrm{T}|\mathcal{H}_1} = \mathrm{E}\{T_{\mathrm{opt}}|\mathcal{H}_1\}$, we have,

$$\mu_{\mathrm{T}|\mathcal{H}_1} = \sum_{i=1}^{M} \frac{\sigma^2\lambda_i}{\sigma^2+\lambda_i}\,\mathrm{E}\{\|\mathbf{w}_i\|^2|\mathcal{H}_1\} = L\sigma^2\sum_{i=1}^{M}\lambda_i, \quad (40)$$

and, setting $\sigma^2_{\mathrm{T}|\mathcal{H}_1} = \mathrm{E}\{T^2_{\mathrm{opt}}|\mathcal{H}_1\} - \mu^2_{\mathrm{T}|\mathcal{H}_1}$,

$$\begin{aligned}\sigma^2_{\mathrm{T}|\mathcal{H}_1} =& \sum_{i=1}^{M}\frac{\sigma^4\lambda_i^2}{(\sigma^2+\lambda_i)^2}\,\mathrm{E}\{\|\mathbf{w}_i\|^4|\mathcal{H}_1\}\\ &+2\sum_{i<k}\frac{\sigma^4\lambda_i\lambda_k}{(\sigma^2+\lambda_i)(\sigma^2+\lambda_k)}\,\mathrm{E}\{\|\mathbf{w}_i\|^2|\mathcal{H}_1\}\,\mathrm{E}\{\|\mathbf{w}_k\|^2|\mathcal{H}_1\}\\ &-L^2\sigma^4\sum_{i=1}^{M}\lambda_i^2 - 2L^2\sigma^4\sum_{i<k}\lambda_i\lambda_k\\ =& L\sigma^4\sum_{i=1}^{M}\lambda_i^2. \end{aligned} \quad (41)$$

Thus, the false-alarm and detection probabilities are calculated as follows,

$$P_{\mathrm{fa}} = \quad P[T_{\mathrm{opt}} > \eta|\mathcal{H}_0] = \quad Q\left(\frac{\eta-\mu_{\mathrm{T}|\mathcal{H}_0}}{\sigma_{\mathrm{T}|\mathcal{H}_0}}\right), \quad (42)$$

$$P_{\mathrm{d}} = \quad P[T_{\mathrm{opt}} > \eta|\mathcal{H}_1] = \quad Q\left(\frac{\eta-\mu_{\mathrm{T}|\mathcal{H}_1}}{\sigma_{\mathrm{T}|\mathcal{H}_1}}\right), \quad (43)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{u^2}{2}\right)\,du$.

B. Performance of T_{R1}

From [33], for a hypothesis test of the form (25), the asymptotic distribution of the Rao test under $\mathcal{H}_0$ and $\mathcal{H}_1$ as $L\to\infty$ is given by,

$$T_{R1} \sim \begin{cases}\chi^2_r & \text{under } \mathcal{H}_0\\ \chi^2_r(\lambda_1) & \text{under } \mathcal{H}_1\end{cases}, \quad (44)$$

where χ^2_r and $\chi^2_r(\lambda_1)$ denote, respectively, central and non-central chi-squared distributions with r degree of freedom and non-centrality parameter λ_1 given by,

$$\lambda_1 = (\boldsymbol{\mu}-\boldsymbol{\mu}_0)^T\mathbf{I}(\boldsymbol{\mu}_0)(\boldsymbol{\mu}-\boldsymbol{\mu}_0). \quad (45)$$

As mentioned in section IV-A, in our case: $\boldsymbol{\mu} = [\varphi_1, \mathrm{Re}\{\varphi_2\},\ldots,\mathrm{Re}\{\varphi_K\},\mathrm{Im}\{\varphi_2\},\ldots,\mathrm{Im}\{\varphi_K\}]^T \in \mathbb{R}^{2K-1}$ and $\boldsymbol{\mu}_0 = \mathbf{0}$. Thus, the decision statistic T_{R1} in (27) is asymptotically distributed as,

$$T_{R1} \sim \begin{cases}\chi^2_{2K-1} & \text{under } \mathcal{H}_0\\ \chi^2_{2K-1}(\lambda_1) & \text{under } \mathcal{H}_1\end{cases}, \quad (46)$$

which, with respect to (68) in Appendix A, the non-centrality parameter λ_1 becomes,

$$\begin{aligned}\lambda_1 =& \boldsymbol{\mu}_r^T\mathbf{I}(\mathbf{0})\boldsymbol{\mu}_r\\ =& \frac{LM}{\sigma^4}\varphi_1^2 + \frac{L}{\sigma^4}\sum_{k=2}^{K}(\mathrm{Re}\{\varphi_k\})^2\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)^2\right)\\ &+\frac{L}{\sigma^4}\sum_{k=2}^{K}(\mathrm{Im}\{\varphi_k\})^2\mathrm{tr}\left((\mathbf{T}_k-\mathbf{T}_k^H)^2\right). \end{aligned} \quad (47)$$

Therefore, the detection and false-alarm probabilities can be evaluated as,

$$P_{\mathrm{fa}} = \frac{\Gamma(\frac{2K-1}{2},\frac{\eta}{2})}{\Gamma(\frac{2K-1}{2})} \quad , \quad P_{\mathrm{d}} = Q_{\frac{2K-1}{2}}\left(\sqrt{\lambda_1},\sqrt{\eta}\right), \quad (48)$$

where $Q_m(\cdot,\cdot)$ denotes the generalized Marcum-Q function and $\Gamma(\cdot,\cdot)$, $\Gamma(\cdot)$ are, respectively, the upper incomplete Gamma and complete Gamma functions.

C. Performance of T_{R2}

From [33], for a hypothesis test of the form (31), the asymptotic distribution of the statistic

$$T_x = (\widehat{\boldsymbol{\mu}}_{r_1}-\boldsymbol{\mu}_{r_0})^T\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}(\widehat{\boldsymbol{\mu}}_{r_1}-\boldsymbol{\mu}_{r_0}), \quad (49)$$

as $L\to\infty$ is given by,

$$T_x \sim \begin{cases}\chi^2_r & \text{under } \mathcal{H}_0\\ \chi^2_r(\lambda_2) & \text{under } \mathcal{H}_1\end{cases}, \quad (50)$$

where the non-centrality parameter λ_2 is given by,

$$\lambda_2 = (\boldsymbol{\mu}_r-\boldsymbol{\mu}_{r_0})^T\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}(\boldsymbol{\mu}_r-\boldsymbol{\mu}_{r_0}), \quad (51)$$

and $\widehat{\boldsymbol{\mu}}_{r_1}$ is the ML estimate of $\boldsymbol{\mu}_r$ under $\mathcal{H}_1$. From [33, App. 6B], as $L\to\infty$, we have,

$$\left.\frac{\partial\mathcal{L}_1(\mathbf{y})}{\partial\boldsymbol{\mu}_r}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} = \left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}(\widehat{\boldsymbol{\mu}}_{r_1}-\boldsymbol{\mu}_{r_0}). \quad (52)$$

By using (52) and also, as in our case $\boldsymbol{\mu}_{r_0} = \mathbf{0}$, we can rewrite (81) in Appendix B as,

$$\begin{aligned}&T_{R2}\\ =&\widehat{\boldsymbol{\mu}}_{r_1}^T\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}[\mathbf{I}^{\dagger}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}\widehat{\boldsymbol{\mu}}_{r_1}\\ \overset{(a)}{=}&\widehat{\boldsymbol{\mu}}_{r_1}^T\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\varphi}\boldsymbol{\varphi}}\right)^{-1}\widehat{\boldsymbol{\mu}}_{r_1}, \end{aligned} \quad (53)$$

where (a) follows from $\mathbf{A}\mathbf{A}^{\dagger}\mathbf{A} = \mathbf{A}$. (53) is the same (49) for $\boldsymbol{\mu}_{r_0} = \mathbf{0}$. Thus, as in our case $\boldsymbol{\mu}_r = [\varphi_1,\mathrm{Re}\{\varphi_2\},\ldots,\mathrm{Re}\{\varphi_K\},\mathrm{Im}\{\varphi_2\},\ldots,\mathrm{Im}\{\varphi_K\}]^T \in \mathbb{R}^{2K-1}$ and $\mu_s = \sigma^2 \in \mathbb{R}$, the asymptotic distribution of the decision statistic in (35) is given by,

$$T_{R2} \sim \begin{cases}\chi^2_{2K-1} & \text{under } \mathcal{H}_0\\ \chi^2_{2K-1}(\lambda_2) & \text{under } \mathcal{H}_1\end{cases}, \quad (54)$$

where, with respect to (80) in Appendix B, the non-centrality parameter λ_2 is obtained as,

$$\begin{aligned}\lambda_2 =& \frac{L}{\sigma^4}\sum_{k=2}^{K}(\mathrm{Re}\{\varphi_k\})^2\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)^2\right)\\ &+\frac{L}{\sigma^4}\sum_{k=2}^{K}(\mathrm{Im}\{\varphi_k\})^2\mathrm{tr}\left((\mathbf{T}_k-\mathbf{T}_k^H)^2\right). \end{aligned} \quad (55)$$

Thus, the detection and false-alarm probabilities can be evaluated as,

$$P_{\mathrm{fa}} = \frac{\Gamma(\frac{2K-1}{2},\frac{\eta}{2})}{\Gamma(\frac{2K-1}{2})} \quad , \quad P_{\mathrm{d}} = Q_{\frac{2K-1}{2}}\left(\sqrt{\lambda_2},\sqrt{\eta}\right). \quad (56)$$

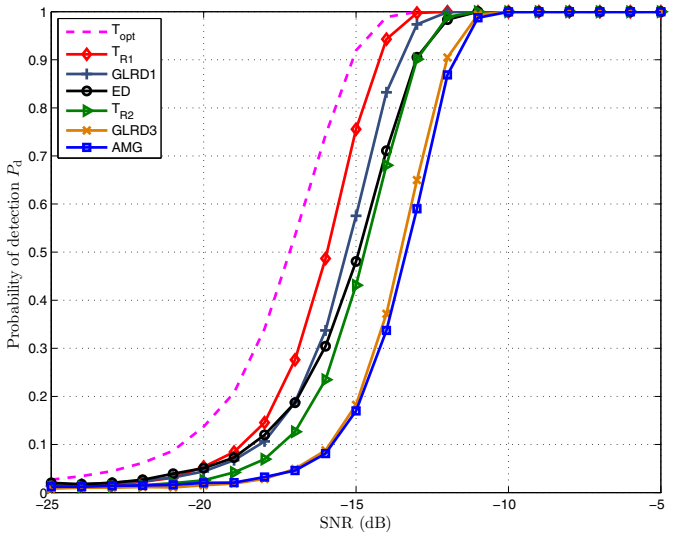

Fig. 1. The probability of detection of different detectors versus SNR for $P_{\rm fa} = 10^{-2}$, $L = 1500$, $M = 4$, $\frac{D}{\lambda_c} = 0.4$ and $\Delta = \pi/8$.

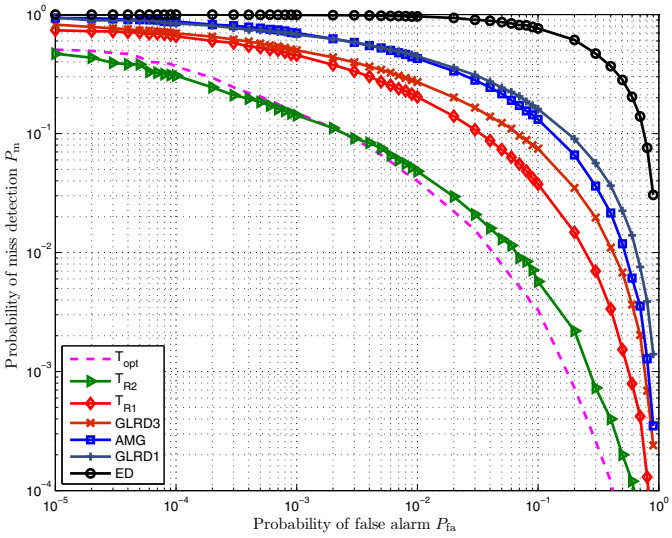

Fig. 3. The complementary ROC of different detectors for average $SNR = -13$ dB, $M = 6$, $L = 1000$, $\frac{D}{\lambda_c} = 0.4$, $\Delta = \pi/8$ and noise variance uncertainty of $\alpha = 0.2$ dB.

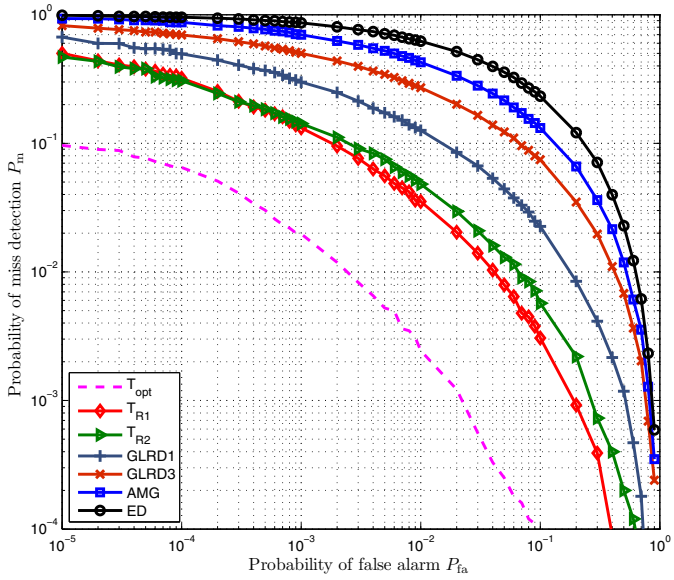

Fig. 2. The complementary ROC of different detectors for average $SNR = -13$ dB, $M = 6$, $L = 1000$, $\frac{D}{\lambda_c} = 0.4$, $\Delta = \pi/8$ and noise variance uncertainty of $\alpha = 0.1$ dB.

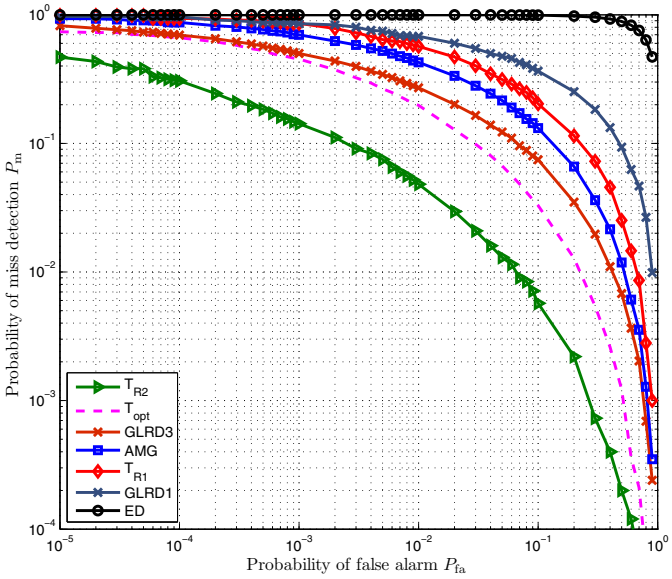

Fig. 4. The complementary ROC of different detectors for average $SNR = -13$ dB, $M = 6$, $L = 1000$, $\frac{D}{\lambda_c} = 0.4$, $\Delta = \pi/8$ and noise variance uncertainty of $\alpha = 0.4$ dB.

VI. Simulation Results

In this section, we provide a comparative simulation-based performance analysis between the proposed detectors and other previously reported detectors used as a benchmark. Specifically, the benchmark detectors are: the GLR Detector-1 (GLRD1) [15, Eqn. (24)], the ED, the GLR Detector-3 (GLRD3) [15, Eqn. (39)] and the Arithmetic to Geometric Mean (AMG) [19, Eqn. (14)] method. Note that: (a) GLRD1 and the ED, similar to T_{R1}, need to know the exact value of the noise variance; (b) GLRD3 and the AMG detector, similar to T_{R2}, are blind detectors. We also compare the simulation-based and analytical-based performance results. The system model used in the simulations is based on the description of Section II, where we assume that the SU is equipped with a linear equispaced antenna array with inter-element spacing D. We assume that the PU system uses the 16-QAM modulation scheme with square-root raised-cosine pulses at 0.25 roll-off and truncated to a length of 25 symbols. Moreover, We assume that the received signal is sampled at the rate of one sample per symbol.

In Figure 1, we depict the probability of detection $P_{\rm d}$ of the optimum detector, T_{R1}, T_{R2} and the four detectors mentioned above versus the average SNR at a false-alarm rate of $P_{\rm fa} = 0.01$, $L = 1500$, $M = 4$, $\frac{D}{\lambda_c} = 0.4$ and $\Delta = \pi/8$. Note that T_{R1}, GLRD1 and the ED assume perfect knowledge of the noise variance. As expected, the performance of the detectors improves with increasing SNR. It can be observed that T_{R1} outperforms the ED and GLRD1. For instance, at

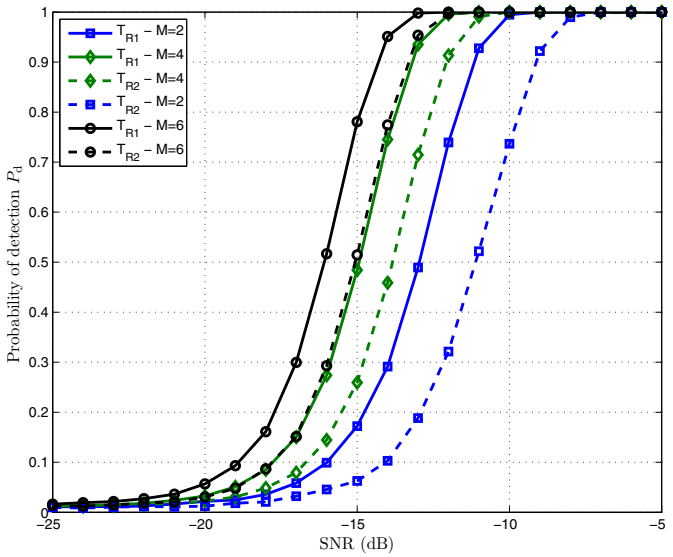

Fig. 5. The probability of detection of the proposed detectors versus SNR for $P_{\rm fa} = 10^{-2}$, $L = 1000$, $\frac{D}{\lambda_c} = 0.4$, $\Delta = \pi/8$ and different values of M.

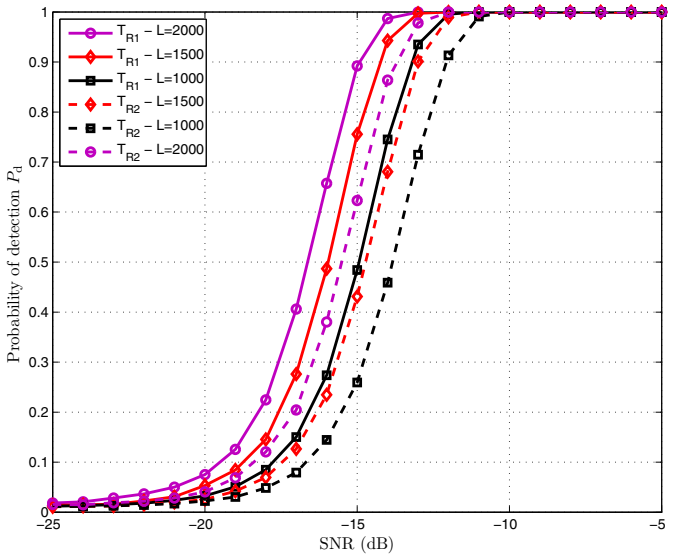

Fig. 6. The probability of detection of the proposed detectors versus SNR for $P_{\rm fa} = 10^{-2}$, $M = 4$, $\frac{D}{\lambda_c} = 0.4$, $\Delta = \pi/8$ and different values of L.

$P_{\rm d} = 0.9$, T_{R1} outperforms GLRD1 by 0.71 dB and the ED by 1.24 dB. The reason T_{R1} outperforms the ED is that it exploits the projections of the off-diagonal elements of the sample correlation matrix, i.e., $\mathrm{tr}(\mathbf{T}_k\mathbf{R})$, in addition to the energy computation of the main diagonal elements. In addition, T_{R1} outperforms GLRD1 because GLRD1 does not consider statistical correlations between the antenna channel gains, but assumes them to be unknown deterministic parameters. We further observe that T_{R2} outperforms the other two blind detectors, i.e., GLRD3 and the AMG detector. In particular, for $P_{\rm d} = 0.9$, the improvement is 1.01 dB over GLRD3 and 1.3 dB over the AMG detector. The reason T_{R2} outperforms GLRD3 is that GLRD3, similarly to GLRD1, ignores statistical correlations between the antenna channel gains. Moreover, T_{R2} shows better performance than the AMG detector since the test statistic of the AMG detector is less informative, e.g., does not exploit the structure of the patterned Hermitian matrix $\mathbf{\Sigma}$ in (9).

Figure 2 shows the complementary ROC (Receiver Operating Characteristics) or the probability of missed detection, i.e., $P_{\rm m}$, versus the probability of false alarm, i.e., $P_{\rm fa}$, for the detectors in Figure 1 under a noise variance mismatch of 0.1 dB for $M = 6$, $L = 1000$, $\frac{D}{\lambda_c} = 0.4$, $\Delta = \pi/8$ and an average SNR of -13 dB. In fact, we assumed that all the detectors overestimated the noise variance by a factor of 0.1 dB and calculated the decision threshold accordingly, but computing the test statistic using data with the actual noise variance. Figure 2 shows that under a noise variance mismatch of 0.1 dB, the optimum detector performs better than T_{R1} and T_{R2}, and that T_{R2} provides sensing performance close to T_{R1}. Also, T_{R1} outperforms the other two blind detectors, i.e., GLRD3 and the AMG detector, while GLRD3 and the AMG detector show better performance than GLRD1 and the ED. Figure 3 shows the same curves depicted in Figure 2 for the same values of M, L, $\frac{D}{\lambda_c}$, Δ and SNR under the noise variance mismatch of 0.2 dB. It can be seen that the performance of the optimum detector, the ED, T_{R1} and GLRD1, degrades significantly in the presence of a greater noise variance mismatch but that T_{R2}, GLRD3 and the AMG detector are robust to the noise variance uncertainty. In fact, under the noise variance mismatch of 0.2 dB, the optimum detector performs similarly to T_{R2}, while T_{R2} outperforms the other detectors. Also, T_{R1} still shows better performance than the other two blind detectors, i.e., GLRD3 and the AMG detector. As can be observed in Figures 2 and 3, T_{R1} is more robust to the noise variance uncertainty than GLRD1 and the ED. Figure 4 shows the same curves depicted in Figures 2 and 3 for the same values of M, L, $\frac{D}{\lambda_c}$, Δ and SNR under the noise variance mismatch of 0.4 dB. Under that noise variance mismatch, the performance of the optimum detector, the ED, T_{R1} and GLRD1 degrades more substantially, so that T_{R2} presents better performance than the optimum detector and GLRD3 and the AMG detector outperform T_{R1}.

In Figure 5, we plot the probability of detection $P_{\rm d}$ of the proposed detectors versus the average SNR at a false-alarm rate of $P_{\rm fa} = 10^{-2}$, $L = 1000$, $\frac{D}{\lambda_c} = 0.4$, $\Delta = \pi/8$ for a varying number of antennas: $M = 2$, 4, 6. As expected, the performance of the proposed detectors improves by increasing the number of antennas. For example, the performance improvements at $P_{\rm d} = 0.9$ from $M = 2$ to $M = 6$ are about 3.15 dB and 4.2 dB for T_{R1} and T_{R2}, respectively. Figure 6 shows the probability of detection $P_{\rm d}$ of the proposed detectors versus the average SNR at a false-alarm rate of $P_{\rm fa} = 10^{-2}$, $M = 4$, $\frac{D}{\lambda_c} = 0.4$, $\Delta = \pi/8$ for a varying number of samples $L = 1000$, 1500, 2000. As can be seen, the performance of the proposed detectors improves accordingly. For instance, at $P_{\rm d} = 0.9$, the performance improves by about 1.75 dB and 1.71 dB for T_{R1} and T_{R2} by increasing the number of samples from $L = 1000$ to $L = 2000$, respectively. However, as expected, the simulation results indicate that increasing the number of antennas, i.e., M, compared to increasing the number of samples, i.e., L, has a more substantial effect on the performance improvement of the proposed detectors.

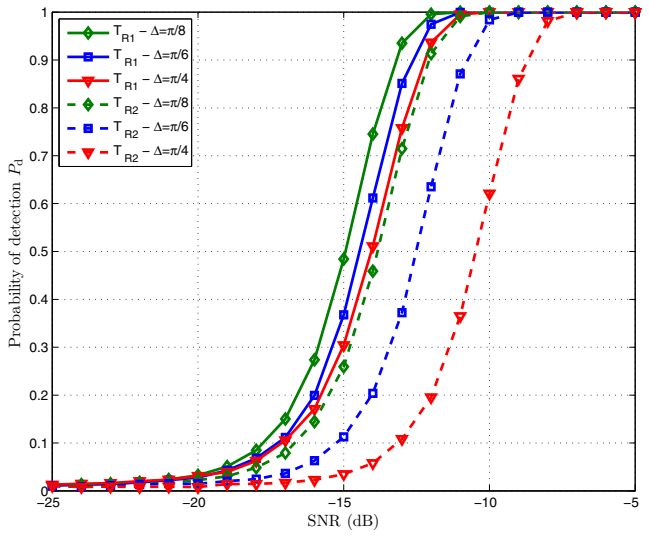

Fig. 7. The probability of detection of the proposed detectors versus SNR for $P_{\rm fa} = 10^{-2}$, $M = 4$, $L = 1000$, $\frac{D}{\lambda_c} = 0.4$ and different values of Δ.

In Figure 7, we depict the probability of detection $P_{\rm d}$ of the proposed detectors versus the average SNR at a false-alarm rate of $P_{\rm fa} = 10^{-2}$, $M = 4$, $L = 1000$, $\frac{D}{\lambda_c} = 0.4$ for different values of Δ: $\Delta = \pi/4$, $\pi/6$, $\pi/8$. Figure 5 shows that by reducing the received signal beamwidth and thus increasing the antenna correlation (as predicted by (7) and (8)), the performance of the proposed detectors improves. In addition, Figure 5 shows that T_{R2} is more sensitive to the antenna correlations than T_{R1}. It is for this reason that, according to (55) and (56), the performance of T_{R2} depends only on the correlation between the antennas, while according to (47) and (48), the performance of T_{R1} depends on both antenna correlations and the energy of the received signal.

In Figures 8(a)–(b), we compare the analytical detection probabilities in Section V with simulation results for detectors T_{opt}, T_{R1} and T_{R2} when $P_{\rm fa} = 0.01$, $L = 1000$, $M = 4$, $\frac{D}{\lambda_c} = 0.4$ and $\Delta = \pi/8$, for two cases: (a) the received signal is sampled at $N_{ss} = 1$ sample per symbol, for which $R_s[m|\mathcal{H}_1] = \mathcal{E}_s\delta[m]$; (b) the received signal is sampled at $N_{ss} = 2$ samples per symbol[1], for which $R_s[m|\mathcal{H}_1] \neq \mathcal{E}_s\delta[m]$. Simulation results are provided for three scenarios in which the channel gain samples are temporally i.i.d. (uncorrelated) and correlated with temporal correlation lengths $L_c = 50$ and $L_c = 100$ samples, spanning 20 and 10 correlation lengths over the 1000-sample detection window, respectively (a Gaussian temporal correlation profile with $\sigma^2 = L_c^2$ is used for the channel gains). We observe that simulation results are very close to the theoretical-based performance for different scenarios even when the received signal is sampled at rate $N_{ss} = 2$. Some loss can be observed in the highest SNR range for $L_c = 100$, which is expected as the detection window only covers 10 correlations lengths. Moreover, as can be seen, the proposed detectors still operate efficiently in the case that $R_s[m|\mathcal{H}_1] \neq \mathcal{E}_s\delta[m]$, and the performance curves are very similar for the same detector

[1]As oversampling by 2 doubles the sampling bandwidth and thus the noise power, the SNR decreases. For the sake of comparison with (a), a corresponding increase by 3 dB has been considered for the primary user E_s/N_0 in (b).

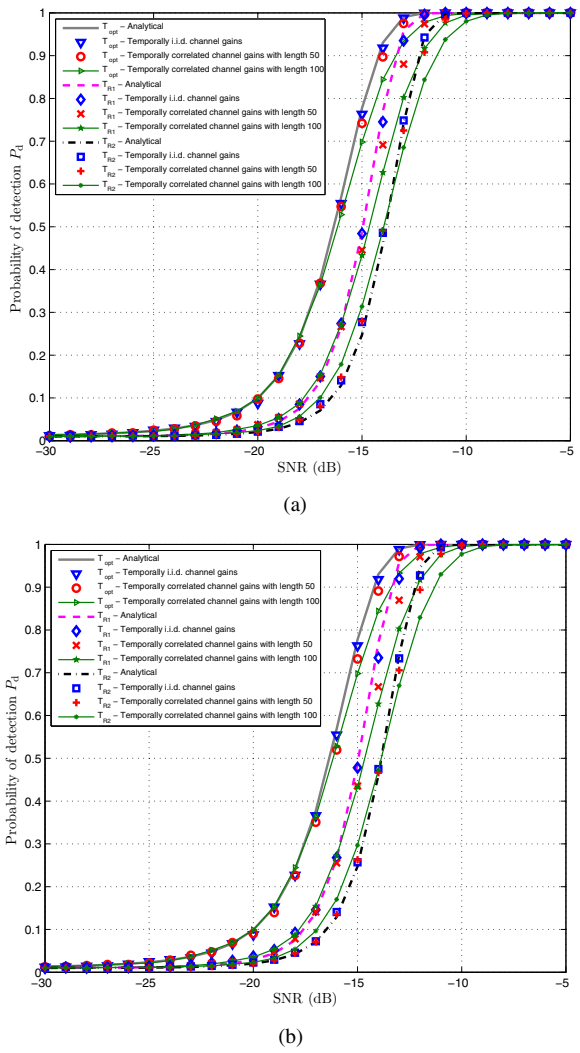

Fig. 8. Comparison between simulation and analytical performance of the proposed detectors, for $P_{\rm fa} = 10^{-2}$, $L = 1000$ and $M = 4$, $\frac{D}{\lambda_c} = 0.4$ and $\Delta = \pi/8$: (a) sampling at the rate of 1 sample/symbol and (b) sampling at the rate of 2 samples/symbol.

compared with case $R_s[m|\mathcal{H}_1] = \mathcal{E}_s\delta[m]$. The roll-off region shifts from left to right as we go from T_{opt} to T_{R1} to T_{R2}, which reflects the corresponding lack of knowledge of signal model parameters.

VII. Conclusion

In this paper, we investigated the spectrum sensing problem for a CR multiantenna receiver which considers the correlation between the received channels at different antennas. We first derived the optimum detector which needs to know the exact value of the antenna correlation coefficients, PU signal power and noise variance. Then, we presented the sub-optimum detectors based on the Rao test for two composite hypothesis scenarios: 1) the antenna correlation coefficients and PU signal power are unknown to the SU; 2) the antenna correlation coefficients, PU signal power and noise variance are unknown to the SU. We also evaluated analytically the performance of the proposed detectors. The simulation results revealed that our proposed detectors outperform several recently-proposed multiantenna detectors for CR.

APPENDIX A
DERIVATION OF T_{R1}

From (24), we have that,

$$\left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \varphi_1}\right|_{\boldsymbol{\mu}=\mathbf{0}} = -\frac{L\mathrm{tr}(\mathbf{T}_1)}{\sigma^2} + \frac{L\mathrm{tr}(\mathbf{T}_1\mathbf{R})}{\sigma^4}, \tag{57}$$

and since $\mathbf{T}_1 = \mathbf{I}_M$, we obtain,

$$\left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \varphi_1}\right|_{\boldsymbol{\mu}=\mathbf{0}} = -\frac{LM}{\sigma^2} + \frac{L\mathrm{tr}(\mathbf{R})}{\sigma^4}. \tag{58}$$

Additionally, differentiating (15) with respect to $\mathrm{Re}\{\varphi_k\}$ and $\mathrm{Im}\{\varphi_k\}$, $2 \le k \le K$, we find,

$$\begin{cases} \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \mathrm{Re}\{\varphi_k\}}\right|_{\boldsymbol{\mu}=\mathbf{0}} = -\frac{L\mathrm{tr}(\mathbf{T}_k+\mathbf{T}_k^H)}{\sigma^2} + \frac{L\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)\mathbf{R}\right)}{\sigma^4}, \\ \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \mathrm{Im}\{\varphi_k\}}\right|_{\boldsymbol{\mu}=\mathbf{0}} = j\left(-\frac{L\mathrm{tr}(\mathbf{T}_k-\mathbf{T}_k^H)}{\sigma^2} + \frac{L\mathrm{tr}\left((\mathbf{T}_k-\mathbf{T}_k^H)\mathbf{R}\right)}{\sigma^4}\right), \end{cases} \tag{59}$$

and with respect to $\mathrm{tr}(\mathbf{T}_k + \mathbf{T}_k^H) = \mathrm{tr}(\mathbf{T}_k - \mathbf{T}_k^H) = 0$, $2 \le k \le K$, we obtain,

$$\begin{cases} \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \mathrm{Re}\{\varphi_k\}}\right|_{\boldsymbol{\varphi}=\mathbf{0}} = \frac{L\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)\mathbf{R}\right)}{\sigma^4}, \\ \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \mathrm{Im}\{\varphi_k\}}\right|_{\boldsymbol{\varphi}=\mathbf{0}} = \frac{jL\mathrm{tr}\left((\mathbf{T}_k-\mathbf{T}_k^H)\mathbf{R}\right)}{\sigma^4}. \end{cases} \tag{60}$$

From [44], the (k,l)-th element of the Fisher information matrix is given by,

$$[\mathbf{I}(\boldsymbol{\mu})]_{k,l} = -\mathrm{E}\left\{\left.\frac{\partial^2 \mathcal{L}_1(\mathbf{y})}{\partial \mu_l \partial \mu_k}\right| \boldsymbol{\mu}\right\}. \tag{61}$$

By calculating $\frac{\partial^2 \mathcal{L}_1(\mathbf{y})}{\partial \mu_l \partial \mu_k}$ for $1 \le k \le l \le 2K-1$ and substituting in (61), we find the different entries of $\mathbf{I}(\boldsymbol{\mu})$ as follows,

$$[\mathbf{I}(\boldsymbol{\mu})]_{1,1} = L\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}\mathbf{T}_1(\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}\mathbf{T}_1\right), \tag{62}$$

for $2 \le k \le K$,

$$[\mathbf{I}(\boldsymbol{\mu})]_{1,k} = L\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}(\mathbf{T}_k+\mathbf{T}_k^H)(\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}\mathbf{T}_1\right), \tag{63}$$

for $K+1 \le k \le 2K-1$,

$$[\mathbf{I}(\boldsymbol{\mu})]_{1,k} = jL\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}(\mathbf{T}_q-\mathbf{T}_q^H)(\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}\mathbf{T}_1\right), \tag{64}$$

for $2 \le k \le l \le K$,

$$[\mathbf{I}(\boldsymbol{\mu})]_{k,l} = L\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}(\mathbf{T}_l+\mathbf{T}_l^H)(\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}(\mathbf{T}_k+\mathbf{T}_k^H)\right), \tag{65}$$

for $K+1 \le k \le l \le 2K-1$,

$$[\mathbf{I}(\boldsymbol{\mu})]_{k,l} = -L\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}(\mathbf{T}_q-\mathbf{T}_q^H)(\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}(\mathbf{T}_p-\mathbf{T}_p^H)\right), \tag{66}$$

for $2 \le k \le K$, $K+1 \le l \le 2K-1$,

$$[\mathbf{I}(\boldsymbol{\mu})]_{k,l} = jL\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}(\mathbf{T}_k+\mathbf{T}_k^H)(\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-1}(\mathbf{T}_q-\mathbf{T}_q^H)\right), \tag{67}$$

with $p = ((k))_K + 1$ and $q = ((l))_K + 1$ which we use the notation $((k))_K$ to denote k modulo K. Also, the Fisher information matrix is a symmetric matrix an thus we have $[\mathbf{I}(\boldsymbol{\mu})]_{k,l} = [\mathbf{I}(\boldsymbol{\mu})]_{l,k}$.

Since $\mathbf{T}_1 = \mathbf{I}_M$, $\mathrm{tr}\left(\mathbf{T}_1(\mathbf{T}_k+\mathbf{T}_k)\right) = \mathrm{tr}\left(\mathbf{T}_1(\mathbf{T}_k-\mathbf{T}_k)\right) = 0$ for $2 \le k \le K$ and also $\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k)(\mathbf{T}_l+\mathbf{T}_l)\right) = 0$ for $k \ne l$, we obtain,

$$[\mathbf{I}(\mathbf{0})]_{k,l} = \begin{cases} \frac{LM}{\sigma^4} & k = l = 1, \\ \frac{L\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)^2\right)}{\sigma^4} & 2 \le k = l \le K, \\ -\frac{L\mathrm{tr}\left((\mathbf{T}_q-\mathbf{T}_q^H)^2\right)}{\sigma^4} & K+1 \le k = l \le 2K-1, \\ 0 & k \ne l. \end{cases} \tag{68}$$

Now, we can obtain the Rao test statistic by substituting (58), (60) and (68) in (26) as,

$$T_{R1} = \sum_{k=2}^{K}\left[\frac{L[\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)\mathbf{R}\right)]^2}{\sigma^4\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)^2\right)} + \frac{L[\mathrm{tr}\left((\mathbf{T}_k-\mathbf{T}_k^H)\mathbf{R}\right)]^2}{\sigma^4\mathrm{tr}\left((\mathbf{T}_k-\mathbf{T}_k^H)^2\right)}\right] + LM\left(\frac{\mathrm{tr}(\mathbf{R})}{M\sigma^2}-1\right)^2 \underset{<\mathcal{H}_0}{\geq\mathcal{H}_1} \eta. \tag{69}$$

Since $\mathbf{R}$ is a Hermitian matrix we can simplify (69) and we obtain the detector specified in (27).

APPENDIX B
DERIVATION OF T_{R2}

From (58) we can evaluate $\left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \varphi_1}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0}$ as,

$$\left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \varphi_1}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} = -\frac{LM}{\widehat{\sigma}_0^2} + \frac{L\mathrm{tr}(\mathbf{R})}{\widehat{\sigma}_0^4}, \tag{70}$$

then, by substituting $\widehat{\sigma}_0^2$ from (28) in (70), we find,

$$\left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \varphi_1}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} = 0. \tag{71}$$

Also, from (60) we can evaluate $\left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \mathrm{Re}\{\varphi_k\}}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0}$ and $\left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \mathrm{Im}\{\varphi_k\}}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0}$, $2 \le k \le K$, as,

$$\begin{cases} \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \mathrm{Re}\{\varphi_k\}}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} = \frac{L\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)\mathbf{R}\right)}{\widehat{\sigma}_0^4}, \\ \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \mathrm{Im}\{\varphi_k\}}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} = \frac{jL\mathrm{tr}\left((\mathbf{T}_k-\mathbf{T}_k^H)\mathbf{R}\right)}{\widehat{\sigma}_0^4}. \end{cases} \tag{72}$$

With respect to (68), we can obtain,

$$\left[\mathbf{I}_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}(\widehat{\boldsymbol{\theta}}_\mathbf{0})\right]_{k,l} = \begin{cases} \frac{LM}{\widehat{\sigma}_0^4} & k = l = 1, \\ \frac{L\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)^2\right)}{\widehat{\sigma}_0^4} & 2 \le k = l \le K, \\ -\frac{L\mathrm{tr}\left((\mathbf{T}_q-\mathbf{T}_q^H)^2\right)}{\widehat{\sigma}_0^4} & K+1 \le k = l \le 2K-1, \\ 0 & k \ne l. \end{cases} \tag{73}$$

From [44], the k-th element of the partition $\mathbf{I}_{\sigma^2\boldsymbol{\mu}_r}(\boldsymbol{\theta})$ from the Fisher information matrix is,

$$\left[\mathbf{I}_{\sigma^2\boldsymbol{\mu}_r}(\boldsymbol{\theta})\right]_k = -\mathrm{E}\left\{\left.\frac{\partial^2 \mathcal{L}_1(\mathbf{y})}{\partial \sigma^2 \partial \mu_{r_k}}\right| \boldsymbol{\mu}_r, \sigma^2\right\}. \tag{74}$$

By calculating $\frac{\partial^2 \mathcal{L}_1(\mathbf{y})}{\partial \sigma^2 \partial \mu_{r_k}}$ and replacing in (74), we have,

$$\left[\mathbf{I}_{\sigma^2 \boldsymbol{\mu}_r}(\boldsymbol{\theta})\right]_k = \begin{cases} L\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-2}\mathbf{T}_1\right) & k=1, \\ L\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-2}(\mathbf{T}_k+\mathbf{T}_k^H)\right) & 2\le k\le K, \\ jL\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-2}(\mathbf{T}_q-\mathbf{T}_q^H)\right) & K+1\le k\le 2K-1. \end{cases} \tag{75}$$

By evaluating $\left[\mathbf{I}_{\sigma^2 \boldsymbol{\mu}_r}(\boldsymbol{\theta})\right]_k$ at $\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0$, we obtain,

$$\left[\mathbf{I}_{\sigma^2 \boldsymbol{\varphi}}(\widehat{\boldsymbol{\theta}}_0)\right]_k = \begin{cases} \frac{LM}{\widehat{\sigma}_0^4} & k=1, \\ 0 & k\neq 1. \end{cases} \tag{76}$$

Also, the partition $\mathbf{I}_{\sigma^2\sigma^2}(\boldsymbol{\theta})$ of the Fisher information matrix is given by,

$$[\mathbf{I}_{\sigma^2\sigma^2}(\boldsymbol{\theta})] = -\mathrm{E}\left\{\frac{\partial^2 \mathcal{L}_1(\mathbf{y})}{\partial\sigma^2\partial\sigma^2}\middle|\boldsymbol{\mu}_r,\sigma^2\right\}. \tag{77}$$

By taking derivative from (30) with respect to σ^2 and replacing in (77), we have,

$$[\mathbf{I}_{\sigma^2\sigma^2}(\boldsymbol{\theta})] = L\mathrm{tr}\left((\boldsymbol{\Sigma}+\sigma^2\mathbf{I}_M)^{-2}\right). \tag{78}$$

By evaluating $[\mathbf{I}_{\sigma^2\sigma^2}(\boldsymbol{\theta})]$ at $\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0$, we obtain,

$$\left[\mathbf{I}_{\sigma^2\sigma^2}(\widehat{\boldsymbol{\theta}}_0)\right] = \frac{LM}{\widehat{\sigma}_0^4}. \tag{79}$$

From (73), (76) and (79), we have,

$$\begin{aligned} &\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1} \\ &= \mathbf{I}_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}(\widehat{\boldsymbol{\theta}}_0) - \mathbf{I}_{\boldsymbol{\mu}_r\sigma^2}(\widehat{\boldsymbol{\theta}}_0)\mathbf{I}^{-1}_{\sigma^2\sigma^2}(\widehat{\boldsymbol{\theta}}_0)\mathbf{I}_{\sigma^2\boldsymbol{\mu}_r}(\widehat{\boldsymbol{\theta}}_0) \\ &= \begin{pmatrix} \frac{LM}{\widehat{\sigma}_0^4} & 0 & \cdots & 0 \\ 0 & \frac{L\mathrm{tr}\left((\mathbf{T}_2+\mathbf{T}_2^H)^2\right)}{\widehat{\sigma}_0^4} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\frac{L\mathrm{tr}\left((\mathbf{T}_K-\mathbf{T}_K^H)^2\right)}{\widehat{\sigma}_0^4} \end{pmatrix} \\ &\quad - \begin{pmatrix} \frac{LM}{\widehat{\sigma}_0^4} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \frac{L\mathrm{tr}\left((\mathbf{T}_2+\mathbf{T}_2^H)^2\right)}{\widehat{\sigma}_0^4} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\frac{L\mathrm{tr}\left((\mathbf{T}_K-\mathbf{T}_K^H)^2\right)}{\widehat{\sigma}_0^4} \end{pmatrix}. \end{aligned} \tag{80}$$

With respect to (80), $\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}$ is a singular matrix and thus the Rao test does not exist for this scenario. So we suggest to use a modified Rao test defined as follows,

$$T_{R2} = \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \boldsymbol{\mu}_r}\right|^T_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0} [\mathbf{I}^{\dagger}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r} \left.\frac{\partial \mathcal{L}_1(\mathbf{y})}{\partial \boldsymbol{\mu}_r}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_0}, \tag{81}$$

with $[\mathbf{I}^{\dagger}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}$ the Moore-Penrose pseudoinverse of $\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}$. The simulation results support this suggested scheme and show that this modified Rao test has a performance close to that of the optimum detector and the ED under perfect knowledge of the noise variance.

The Moore-Penrose pseudoinverse of a singular diagonal matrix is obtained by taking the reciprocal of each non-zero element on the diagonal, leaving the zeros in place, and transposing the resulting matrix. Thus, the Moore-Penrose pseudoinverse of $\left([\mathbf{I}^{-1}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r}\right)^{-1}$ becomes,

$$\begin{aligned} &[\mathbf{I}^{\dagger}(\widehat{\boldsymbol{\theta}}_0)]_{\boldsymbol{\mu}_r\boldsymbol{\mu}_r} \\ &= \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \frac{\widehat{\sigma}_0^4}{L\mathrm{tr}\left((\mathbf{T}_2+\mathbf{T}_2^H)^2\right)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\frac{\widehat{\sigma}_0^4}{L\mathrm{tr}\left((\mathbf{T}_K-\mathbf{T}_K^H)^2\right)} \end{pmatrix}. \end{aligned} \tag{82}$$

Substituting (71), (72) and (82) in (81), the following statistic results for the modified Rao test,

$$\begin{aligned} &T_{R2} \\ &= \sum_{k=2}^{K}\left[\frac{L[\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)\mathbf{R}\right)]^2}{\widehat{\sigma}_0^4\mathrm{tr}\left((\mathbf{T}_k+\mathbf{T}_k^H)^2\right)} + \frac{L[\mathrm{tr}\left((\mathbf{T}_k-\mathbf{T}_k^H)\mathbf{R}\right)]^2}{\widehat{\sigma}_0^4\mathrm{tr}\left((\mathbf{T}_k-\mathbf{T}_k^H)^2\right)}\right] \\ &\underset{<}{\overset{>}{}}\begin{matrix}\mathcal{H}_1\\ \mathcal{H}_0\end{matrix} \eta. \end{aligned} \tag{83}$$

Since $\mathbf{R}$ is a Hermitian matrix we can simplify (83) and we obtain the detector specified in (35).

REFERENCES

[1] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surv. Tutor.*, vol. 11, no. 1, pp. 116–130, 2009.

[2] E. Axell, G. Leus, E. Larsson, and H. Poor, "Spectrum sensing for cognitive radio: state-of-the-art and recent advances," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 101-116, May 2012.

[3] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proc. IEEE*, vol. 55, no. 4, pp. 523-531, Apr. 1967.

[4] F. F. Digham, M. S. Alouini, and M. K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 21-24, Jan. 2007.

[5] R. Tandra and A. Sahai, "SNR walls for signal detection," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 4-17, Feb. 2008.

[6] A. Taherpour, Y. Norouzi, M. Nasiri-Kenari, A. Jamshidi, and Z. Zeinalpour-Yazdi, "Asymptotically optimum detection of primary user in cognitive radio networks," *IET Commun.*, vol. 1, pp. 1138-1145, Dec. 2007.

[7] A. Singh, M. R. Bhatnagar, and R. K. Mallik, "Optimization of cooperative spectrum sensing with an improved energy detector over imperfect reporting channels," in *Proc. 2011 IEEE Veh. Tech. Conf. – Fall*, pp. 1-5.

[8] Q. Chen, M. Motani, W.-C. Wong, and A. Nallanathan, "Cooperative spectrum sensing strategies for cognitive radio mesh networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 56-67, Feb. 2011.

[9] G. Taricco, "Optimization of linear cooperative spectrum sensing for cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 77-86, Feb. 2011.

[10] Y. Zou, Y.-D. Yao, and B. Zheng, "A selective-relay based cooperative spectrum sensing scheme without dedicated reporting channels in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1188-1198, Apr. 2011.

[11] Y. Zou, Y.-D. Yao, and B. Zheng, "Cooperative relay techniques for cognitive radio systems: spectrum sensing and secondary user transmissions," *IEEE Commun. Mag.*, vol. 50, no. 4, pp. 98-103, Apr. 2012.

[12] L. Shen, H. Wang, W. Zhang, and Z. Zhao, "Multiple antennas assisted blind spectrum sensing in cognitive radio channels," *IEEE Commun. Lett.*, vol. 16, no. 1, pp. 92-94, Jan. 2012.

[13] E. Axell and E. Larsson, "Optimal and sub-optimal spectrum sensing of OFDM signals in known and unknown noise variance," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 290-304, Feb. 2011.

[14] H. Nguyen, E. De Carvalho, and R. Prasad, "Spectrum sensing for cognitive radio based on multiple antennas," in *Proc. 2012 IEEE Veh. Tech. Conf. – Spring*, pp. 1-5.
[15] A. Taherpour, M. Nasiri-Kenari, and S. Gazor, "Multiple antenna spectrum sensing in cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 814-823, Feb. 2010.
[16] R. López-Valcarce, G. Vazquez-Vilar, and J. Sala, "Multiantenna spectrum sensing for cognitive radio: overcoming noise uncertainty," in *Proc. 2010 Int. Workshop on Cognitive Inform. Process.*, pp. 310-315.
[17] P. Bianchi, M. Debbah, M. Maïda, and J. Najim, "Performance of statistical tests for single-source detection using random matrix theory," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2400-2419, Apr. 2011.
[18] P. Wang, J. Fang, N. Han, and H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1791-1800, May 2010.
[19] R. Zhang, T. Lim, Y. Liang, and Y. Zeng, "Multi-antenna based spectrum sensing for cognitive radios: a GLRT approach," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 84-88, Jan. 2010.
[20] D. Ramírez, G. Vazquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Detection of rank-p signals in cognitive radio networks with uncalibrated multiple antennas," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3764-3774, Aug. 2011.
[21] J. Salz and J. Winters, "Effect of fading correlation on adaptive arrays in digital mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 4, pp. 1049-1057, Nov. 1994.
[22] G. Durgin and T. Rappaport, "Effects of multipath angular spread on the spatial cross-correlation of received voltage envelopes," in *Proc. 1999 IEEE Veh. Tech. Conf. – Spring*, pp. 996-1000.
[23] IEEE 802.22, Working Group on Wireless Regional Area Networks (WRAN). Available: http://grouper.ieee.org/groups/802/22/.
[24] C. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. Shellhammer, and W. Caldwell, "IEEE 802.22: the first cognitive radio wireless regional area network standard," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 130-138, Jan. 2009.
[25] S. Kim, J. Lee, H. Wang, and D. Hong, "Sensing performance of energy detector with correlated multiple antennas," *IEEE Signal Process. Lett.*, vol. 16, no. 8, pp. 671-674, Aug. 2009.
[26] V. R. S. Banjade, N. Rajatheva, and C. Tellambura, "Performance analysis of energy detection with multiple correlated antenna cognitive radio in Nakagami-m fading," *IEEE Commun. Lett.*, vol. 16, no. 4, pp. 502-505, Apr. 2012.
[27] L. Luo, P. Zhang, G. Zhang, and J. Qin, "Spectrum sensing for cognitive radio networks with correlated multiple antennas," *Electron. Lett.*, vol. 47, no. 23, pp. 1297-1298, Nov. 2011.
[28] K. Hassan, R. Gautier, I. Dayoub, E. Radoi, and M. Berbineau, "Predicted eigenvalue threshold based spectrum sensing with correlated multiple-antennas," in *Proc. 2012 IEEE Veh. Tech. Conf. – Spring*, pp. 1-5.
[29] M. Orooji, R. Soosahabi, and M. Naraghi-Pour, "Blind spectrum sensing using antenna arrays and path correlation," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 3758-3767, Oct. 2011.
[30] D. Romero, J. Vía, R. López-Valcarce, and I. Santamaría, "Detection of Gaussian signals in unknown time-varying channels," in *Proc. 2012 IEEE Statistical Signal Process. Workshop*, pp. 916-919.
[31] D. Ramírez, J. Iscar, J. Vía, and L. L. Scharf, "The locally most powerful invariant test for detecting a rank-p Gaussian signal in white noise," in *Proc. 2012 IEEE Sensor Array and Multichannel Signal Process. Workshop*, pp. 501-504.
[32] H. Van Trees, *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*. John Wiley and Sons Inc., 2002.
[33] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory.* Prentice Hall, 1998.
[34] G. Vazquez-Vilar, R. López-Valcarce, and J. Sala, "Multiantenna spectrum sensing exploiting spectral a priori information," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 4345-4355, Dec. 2005.
[35] J. Sala, G. Vazquez-Vilar, and R. López-Valcarce, "Multiantenna GLR detection of rank-one signals with known power spectrum in white noise with unknown spatial correlation," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3065-3078, Jun. 2012.
[36] J. Villares Piera, "Sample covariance based parameter estimation for digital communications," Ph.D. dissertation, Tech. University of Catalonia, 2005.
[37] J. Villares and G. Vázquez, "The Gaussian assumption in second-order estimation problems in digital communications," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4994-5002, Oct. 2007.
[38] T. Kollo and K. Ruul, "Approximations to the distribution of the sample correlation matrix," *J. Multivar. Anal.*, vol. 85, no. 2, pp. 318-334, May 2003.
[39] H. Neudecker and A. Wesselman, "The asymptotic variance matrix of the sample correlation matrix," *Linear Algebra Applicat.*, vol. 127, pp. 589-599, 1990.
[40] M. Browne and A. Shapiro, "The asymptotic covariance matrix of sample correlation coefficients under general conditions," *Linear Algebra Applicat.*, vol. 82, pp. 169-176, Oct. 1986.
[41] A. Manikas, *Differential geometry in array processing.* Imperial College Press, 2004.
[42] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra.* Society for Industrial and Applied Mathematics, 2000.
[43] A. Papoulis, *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, 2002.
[44] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory.* Prentice Hall, 1993.

Saeid Sedighi was born in Tehran, Iran, in 1987. He received the B.Sc degree in Electrical Engineering from the Islamic Azad University, Tehran, Iran, in 2010 and the M.Sc. degree in Telecommunications Engineering from Imam Khomeini International University (IKIU), Qazvin, Iran, in 2012. Since Otober 2012, he has served as a research assistant in the Signal Processing and Information Theory (SPIT) lab at the Electrical Engineering Department, Imam Khomeini International University (IKIU). His research interests are cognitive radio, statistical signal processing, detection and estimation theory, and cooperative communications.

Abbas Taherpour (S07-M09) received the B.Sc. degree in electrical engineering in 2002 from Sharif University of Technology, Tehran, Iran, and the M.Sc. degree in communication systems engineering in 2004, from Tehran University, Tehran, Iran. He also received his Ph.D. in communication systems engineering from Sharif University of Technology, Tehran, Iran, in 2009. He has been also a visiting scholar at Queens University, ON, Canada form June 2007 until June 2008. He is currently an assistant professor with the department of electrical engineering, Imam Khomeini International University (IKIU), Qazvin, Iran. His research interests are cognitive radio, statistical signal processing and wireless communications.

Josep Sala-Alvarez (S'86 - SM'11) was born in Barcelona, Spain, in 1967. He received the M.Sc. and Ph.D. degrees in telecommunications engineering from the Technical University of Catalonia - Barcelona Tech (UPC), Barcelona, Spain, in 1991 and 1995, respectively. During 1992, he worked at the European Space Operations Centre (ESOC) of the European Space Agency (ESA), Darmstadt, Germany, in the area of software engineering for telemetry processing. From 1993 to late 1994, he held a grant from the Generalitat de Catalunya in support of the Ph.D. degree at the Department of Signal Theory and Communications, UPC. In 1994, he joined this department as Assistant Professor and was promoted to Associate Professor in 1997. He has participated in space-related communication projects for ESA and in wireless communications projects at the national/European level with industry and institutions. His current research interests are in the field of signal processing, communications, and information theory. Dr. Sala is recipient of the IEEE Signal Processing Society Best (Senior) PaperAward 2003, the International Symposium on Turbo-Codes and Applications (ISTC03) Best Poster Paper Award, and the Best Ph.D. Thesis in Telecommunications National Award (Spain, 1995). Dr. Sala is an IEEE Senior Member.